

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:36:02 ; Search time 157.61 Seconds
(without alignments)
7900.362 Million cell updates/sec

Title: US-09-715-909-1

Perfect score: 5498
Sequence: 1-cataatacacaataagagga.....aaaaaaaaaacctcag 5498

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2.6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2.6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2.6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2.6/ptodata/2/ina/PCITUS.COMB.seq:*
- 6: /cgn2.6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1954.6	35.6	5577	1	US-08-326-117B-1
2	1954.6	35.6	5577	3	US-08-982-129-1
3	57.2	1.0	790	6	PCT-US92-05532-1
4	55	1.0	1910	5	PCT-US92-05532-1
5	55	1.0	2261	1	US-08-272-882D-1
6	53.6	1.0	789	6	5219739-8
7	53.2	1.0	1294	3	US-09-025-691-2
8	51	0.9	3377	6	5198345-16
9	50.8	0.9	4104	1	US-07-998-003A-94
10	50.8	0.9	4104	1	US-08-453-274B-94
11	50.8	0.9	4104	1	US-08-453-695A-94
12	50.8	0.9	4104	1	US-08-268-161A-94
13	50.8	0.9	4104	2	US-08-453-702A-94
14	50.8	0.9	4104	4	US-09-099-639-94
15	50.8	0.9	4104	5	PCT-US93-12588-94
16	50.8	0.9	4104	5	PCT-US95-08071-94
17	50.8	0.9	4650	1	US-07-998-003A-102
18	50.8	0.9	4650	1	US-08-453-274B-102
19	50.8	0.9	4650	1	US-08-453-695A-102
20	50.8	0.9	4650	1	US-08-268-161A-102
21	50.8	0.9	4650	2	US-08-453-702A-102
22	50.8	0.9	4650	2	US-09-099-639-102
23	50.8	0.9	4650	5	PCT-US93-12588-102
24	50.8	0.9	4650	5	PCT-US95-08071-102
25	49.2	0.9	1288	1	US-08-440-856A-9
26	49.2	0.9	4972	3	US-09-035-648-17
27	49.2	0.9	4972	4	US-09-001-951-17

28	48.4	0.9	1302	1	US-08-525-697-1	Sequence 1, Appli
29	46.8	0.9	1656	1	US-08-741-881-1	Sequence 1, Appli
30	46.8	0.9	1656	1	US-08-739-158-1	Sequence 1, Appli
31	46.8	0.9	1656	2	US-08-739-167-1	Sequence 1, Appli
32	46.8	0.9	1656	3	US-08-404-796-1	Sequence 1, Appli
33	46.8	0.9	1656	3	US-08-931-869-1	Sequence 1, Appli
34	46	0.8	1272	2	US-09-068-109-1	Sequence 1, Appli
35	46	0.8	2712	3	US-09-025-691-4	Sequence 1, Appli
36	45.8	0.8	2007	3	US-08-747-221B-36	Sequence 36, Appli
37	45.8	0.8	2007	3	US-08-747-221B-38	Sequence 36, Appli
38	45.8	0.8	2007	4	US-09-005-051-36	Sequence 38, Appli
c 39	45.8	0.8	2007	4	US-09-005-051-38	Sequence 38, Appli
40	44.6	0.8	1035	4	US-09-105-390-49	Sequence 49, Appli
41	44.4	0.8	2230	4	US-09-197-649-7	Sequence 1, Appli
42	44.2	0.8	390	4	US-09-197-649-7	Sequence 7, Appli
43	44.2	0.8	924	4	US-09-105-390-33	Sequence 33, Appli
44	44	0.8	1431	4	US-09-078-173A-7	Sequence 7, Appli
45	44	0.8	1443	4	US-09-078-173A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-326-117B-1
Sequence 1 Application US/08326117B
Patent No 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5577 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 197..4780
US-08-326-117B-1

Query Match 35.6%, Score 1954.6; DB 1; Length 5577;
Best Local Similarity 63.9%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 1769; Indels 62; Gaps 17;

QY	1211	cgaaagacttgaagagagacttttcaaccctccaccatcogtttgtcttaacaatctctgaagc	1270
Db	1261	CGACCACTGCAACGAAAGAGTGTGTTCCACTTACGATCGTGGCTTTACAAAATATGAGGA	1320
QY	1271	tagctttgcaacgagagcccaacttttccatcatcgttcacgaactcaatpatcagaagac	1330
Db	1321	GGCCTTCTCCACATCAACAACAAACGTGGTCACTCACTATTGTGACAGACATCAGACCAAGAAC	1380
QY	1331	cgagcgcgtgataaagaataglatatgatalcatatgaaagaaactccaatgaactctaa	1390
Db	1381	TGAACCTATACNCAAGGAATATGCACTGGCAATCATGAGAGAGACGCCCTGACCTCAA	1440
QY	1391	cttcaatgaagaactttgattccatgcatgaatttgggtgaaagacttaataacagf	1450
Db	1441	CTTCGATTAAGAAATTCGGATTCTCATGTATAGATTATGGTCAAAACCTCATGTAACGGT	1500
QY	1451	ggaacttgagagacttgttcccgccagggggggtccgacattcaactcgcgcggggag	1510
Db	1501	GGGTCTAGAGAGCGTTGACCCCTCAAGCGCTGAGCGCATTTCTACTACGCGCTGAAGT	1560
QY	1511	cgagtcaccagagggagagacttcatcatatgagccataaacaacaccatgtctgattacga	1570
Db	1561	CGGCTACCGAGGACAGACACTTCATCATGAGGAGACCCCTCAATCACTCCATGCGATTACGA	1620
QY	1571	agatgtcatltttcagaacatcatcaatlaagfccaagcagltgacatgaaacagctag	1630
Db	1621	AGTGCCAGAGTTTCAGATATTACGATTCGGGTGTTAGCGCACCCACAAACGACACGAG	1680
QY	1631	ccaagttggcgaagcgctggtgtacgtgaaccttgataactcggagagagaacttccat	1690
Db	1681	GCACGTGGCGCTCCGGTGGTTACATTTGACCTCATATTTGAAACATGACATGACGCCAT	1740
QY	1691	cttcgaagagagcagctacatcctcgcgcgttgtttaaagagaacgctgcgcgcgcctccggt	1750
Db	1741	CTTCCAAACACCGCGTGCAGACCGCTCACCTTGAGAGAGAGCTAAGACGAGGGGTCTTGCT	1800
QY	1751	ggccacggtgtctgcgccttcgacagagacatcgacgaagctagtagtcatcatatgattgg	1810
Db	1801	CGCCAAAGCGGTGCACACGACAGACATCGGGGAGTGCCTCGAGAGACTTCTTATTGGG	1860
QY	1811	caagcgttttactactcgtgtcatatagatgaatcaacgggaaagcttctcgtgagcatgga	1870
Db	1861	TAAACCTGTATTAATTCCTCGACCATTCGACAAACTCACCGCGCACATCCGCGTTCAGCTAA	1920
QY	1871	cgaigtcttgactacacacgacagaacacttatattgtcaagttgcgcttgacgaatc	1930
Db	1921	CGACTTCCTTCACTACCTACCTCGAAGAAATGTAATTTTGTGAGGTGCGACGCTACAGACAC	1980
QY	1931	tttgggcgaagcggccacacaacacagltgacacaccagcttgtatgaagtgaagatgt	1990
Db	1981	GCTGGGCGCA --- ACCCTTCCACACGGGCGACGTCACAGCTGGTCATGACATTAATGTGACAT	2037
QY	1991	caacaacactcctccaccctacgctgtgcgcgcgttcgaactccaagcgttcgagggagaagc	2050
Db	2038	CAACAACACGCGCACCCACCTTACGGCTGCTCCTCGAAGCGAGTCCCAAGTGGAGGAAAGT	2097
QY	2051	tcccgaagatacagatatcccggaataatcactgtctcagaaacccggagacacagcagccta	2110
Db	2098	GGCTATATGGCCAGTCATCACCCACAGAGATTACGCCACACCGACCCGACACACAGGCCGA	2157
QY	2111	cctgtgatttcagatcagacttggaacttcaacttgggcacacaaagcagggcagaagacata	2170
Db	2158	TCTGGCTTTCGAGATTAACACTGGGACACTCTTTGGCCACCAAGAGGCCGCCAAGGCTAA	2217
QY	2171	cctactatgaatacgtcgggtglatagttatgaagaagatalaaccacacagggagcaacg	2230
Db	2218	CCCCGACGAGTTTAGAAATTGGCTGGAATATCGAGACCATCTCCGCCGAGATTAAACAAACG	2277
QY	2231	gggttcgcgaatcgggcgcctcgtgtgtcgaagagatcccggaacaacgtcacacatcagctt	2290
Db	2278	GGGACTGGGTATTCGGCCCGGTTGTAGCCGGCGCAAAATTCAGACCAACATGTAACATATGACTA	2337
QY	2291	ggaagaaattcgaatagctttactccacgctcgcgttgagggagaccccaacacgtgcatcgg	2350

Db	2338	CGAGGAGTTTGAAGGTCCTCTCCCTCCACAGTGAAGGCTGCGTACCTTAAACCGCTACGG	2397
OY	2351	agatgacacgaatgaaagcgacgcttccacgatacacaataatcagatcgtaaagacacgccc	2410
Db	2398	AGAGACACTACGACGAATTCGATCGTCAACAAATATATATGATATGAAACGACACGGCC	2455
OY	2411	gacttcgcgaacgagcgacgcctgacgaacgacgatgacgctgacgagctggcgcgacgag	2470
Db	2458	GGTGTGGGTGGAGGGGACCTCTGGAGCGAGAACTTCCAGATCCCGGAGATGTCCGGCGGG	2517
OY	2471	caacgctaagcgcctccgtgctgcgacacgacatcagacgcccgcgtctaaaccaagtgcg	2530
Db	2518	GCTGTGGTGGGCTCCGTGGCGGGGACGACATCGACGGACCGCTCTTACAAACCAAGTGG	2577
OY	2531	ctacacatacaaacctagaacaacaacatcccgagagattagtgaaagtgtactcaaac	2590
Db	2578	ATACGACCAATTTTCCCTCGTGAAGACACAGATTAAGACACTGATATATATGACTTCTCAC	2637
OY	2591	tggtlcaaatgtgaagtgatgcgaacgaagcgatcgatgacgaacacccctgacgttcta	2650
Db	2638	GGGTCAAAATTTCCGTGAACAACAACCGCGCCATCGACGGGAGATCTCCACCGTTCCA	2697
OY	2651	cttgctacacacccgtatcgctgacgacgagtgctccctggaaaacgcgagaattgcc	2710
Db	2698	CCTCTACTATACACTGGTGCCTATGTGACCCATCTCGACAGAAATCCMGCAATTTGCC	2757
OY	2711	tccagattccaaactcactctgaaatgccagatcgcgaatagaaatcatctgcacaaa	2770
Db	2758	CCCTGACCCGACCTATTGTGGGAACCGAAGGAATATCACAAATCCACATACCGACACGA	2817
OY	2771	caacaaagtgcctcgagccgctcacatgaaagtccaacacgaacggtgactctggagaa	2830
Db	2818	CAACAGAGTCCCGCAGCGGAACAGACTAGTTTCGATACCGTCTGATATATTACGAGA	2877
OY	2831	tgccaacgagcgagcgaaaggttgctcaagctgtatctccacgacgctgaaagaagagtc	2890
Db	2878	CGCAACCCACTTAAACGAGGTGGTCACCTGATAGGCCAATGATCTTGACACGAGACGAAT	2937
OY	2891	gtacacacacgttaacatacacaatgtaactttgcgtgaaaccccgacgtcgagattcct	2950
Db	2938	ATACACACAGGTGAGCTACGTCAATTAATGACAGGAACCTCGACTATGAACTTCTT	2997
OY	2951	cgaggtggaacctggaacactgctgcgcttgaagtgataccgg-----ggagaca	3001
Db	2998	CTCCGTGAACCGAGAGACCGCGCTGCTGATACGTCGACTATGAGACCCACAGGATAGGGCGA	3057
OY	3002	aaattggaacgagatgagagataagcttaacatactactactcttgaattcatcgtataa	3061
Db	3058	GGTGTGACCGCTGATGATGTATTAACCAAGCACCGTATCTTCTTCAACCTCATCGACAA	3117
OY	3062	ctctctttctgatatgtaagtagtagagaacacgaagcgagttgaaatattgtcgtct	3121
Db	3118	CTTATGAGGGGAAGGAGAAAGTACAGAAATCAGAACGACACAGAAATCTTCGTATATT	3177
OY	3122	atggatgtgaacacacaacgctctcctgagtgatgcatctgcatctgtaactcgcgtttagtc	3181
Db	3178	GTTGGAATGTGAATGACAAATGCTCTCTGAATTTGCCACCGCGACGCAACTCTGTGACAT	3237
OY	3182	ttcggaaggaagcagtgctgagtgctccgtgtaactccagaatctacgacacggagacaga	3241
Db	3238	ATTCGAAACACTTAAAGCAGGGCGCTCGCTGTGAAACACATATCTTCCGCCGCGACCGGA	3297
OY	3242	tgaaacgaacacggaacaacacgctgcttggttagaagaaaccccgagcccaagatcagca	3301
Db	3298	CGAGCCGACACACACACACTCCAGGGTGGCTTACGAGATCTCGAAACTCT--AGACGGA	3354
OY	3302	ccgaagacatcgaggtgacggaatctctcacaatgatactgatactgaaacaaacacgtggga	3361
Db	3355	GGCGGACATCGAAGTGGCGGAGCTGTTTGATGATATACAGATCGGGAACGTCCAGGGAGA	3414
OY	3362	acttggaacgcgataatgagacttgagggaggtatctggggcaacttaacgaatatcatctgagc	3421

[illegible]

[illegible]

	APPLICATION NUMBER:	US/08/982,129
	FILING DATE:	
	CLASSIFICATION:	
	PRIOR APPLICATION DATA:	US 08/326,117
	APPLICATION NUMBER:	19-OCT-1994
	FILING DATE:	19-OCT-1994
	ATTORNEY/AGENT INFORMATION:	
	NAME:	MILLMAN, ROBERT A.
	REGISTRATION NUMBER:	36,217
	REFERENCE/DOCKET NUMBER:	7112-0037.00
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE:	(202) 887-1500
	TELEFAX:	(202) 887-0763
	TELEX:	90-4030
	INFORMATION FOR SEQ ID NO: 1:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH:	5577 base pairs
	TYPE:	nucleic acid
	STRANDEDNESS:	double
	TOPOLOGY:	linear
	MOLECULE TYPE:	cDNA
	FEATURE:	
	NAME/KEY:	CDS
	LOCATION:	197..4780
	US-08-982-129-1	
	Query Match	35.6%; Score 1954.6; DB 3; Length 5577;
	Best Local Similarity	63.9%; Pred. No. 0;
	Matches 3246; Conservative	0; Mismatches 1769; Indels 62; Gaps
Dn	187 ATATGTGAGATGGCAGTGTGACGCCGAATCGCTTCCTGCTGTTTATAGCGCC	246
Oy	152 atcatgcaaaatgggggttgagaggtcttccacgaagtactgltccttagcttc	211
Dn	212 tgcgcgaactagccaaccaagcatgttcgatcatatgacaataccaaggccgaagatccc	271
Oy	212 tgcgcgaactagccaaccaagcatgttcgatcatatgacaataccaaggccgaagatccc	271
Dn	247 TGCAGTTTTACTCTAAGAAGAAATGGGTTATGATGACC GCCATCCCAAGGCTAC CAGACC	306
Oy	272 ggaac---tgcgcgctaattgatlcagaagaaaatacctggagtgaagacccttaaac	328
Dn	307 GGATAATTGTCAGTACTAATAATTTGGAAGGCACAGACATGGAGTGTCAGAGCCCTCGTCCC	366
Oy	329 cgcgccgaaccggagaaagtatgtatgtgagaact-----tcttacggataaat	379
Dn	367 CGCCCCGGAGGGGATGACCTGTGCATGGACGCCCTPACCGTGAATAACACCAACTCGG	426
Oy	380 gatcagaatcatatcatatggagaaatcgaaagagagcagatcatatctgaaagcttaa	439
Dn	427 CACGAGGTCTATCATCATGATGATGAAGAGATAGAAAGACAATCACCATGCCATFACTTA	486
Oy	440 ctatcaagggtccaacaagccggtgtgtctgcgatatgtcgaagcccaagcccaagct	499
Dn	487 TTAT-AACGGACATCAACTCCGTTCATTGAACTGCCATTTTATCCGGTTCTACAAATC	545
Oy	500 ggagcccttagttctggacaagaatgaagcagaagcccaatggagccctgttatctgcgaag	559
Dn	546 TGCTGATGCCGCTATCATCAGAGATTTGACAACGGGATGCAATC--TCATATCAACGCAAG	603
Oy	560 acaagactacgaagacgaacacatctgacgaagctatgtgttctcaatccaaatggaggtga	619
Dn	604 ACACATTCACAGATTGCCCGCGCATGACGACGATGATTTCAATGTCGGCTGGACGGCCA	663
Oy	620 ataacagccgctactgttgtggtcgtgagatagtccaatlgaaagaaatccggccatcct	679
Dn	664 GTTCGTGTGTCAGACCGGTGTCTCGCTATGCTAATATGATGATACACCGCCCATCAT	723
Oy	680 gcaaagtgtatcgagccctgtgcgttaattctcagaacaatctgagagctgaagctgacctgt	739
Dn	724 ACAAAACTTCAGAGCTTGCCCGGTTCTCTGAACCTGGGAGGCAAGGTTGACAGATGCAC	783
Oy	740 gtaccaaagtgtcacagcggaagtgtaaatatgaacacccgcgtcaatgaagcttcggtgtcga	799

OY	2951	cgaggtggacccctggacaactggtctgcctctgaagttgcattaccggy-----ggagca	3001
Db	2998	CTCCGCTGAACCGGAGAACCGGCGCTGGTGTACGTGGACATATGAGCAACGAGGTATGGGGCA	3057
OY	3002	aaatctggaccgcgcatggggatgagcctcacaatactactctgtttaaatttcattcgcaca	3061
Db	3058	GGTGGCTGACCGGATGGTATGATTAACCAACGACCGGATCTCTTCAACCTCATGACAA	3117
OY	3062	cttcctttctgattggttgacggttagagaaacaaagacgaagttgaaatattgtcgtctc	3121
Db	3118	CTTCATGGGGGAGGAGAAAGTATACGAAGATACGAACGACAGAAATTCCTGTTATCTT	3177
OY	3122	attgagatgttaacagcaaacgcctcttgaatgtccattgtccgtgaatgcacccggtttgagtt	3181
Db	3178	GTTGGATGTGAATGACAAATGCTCTCTTAATTTGCCACCGCGGACGAACTCTCTTGGACATAT	3237
OY	3182	ttccgaagagacagttgtctgtgtgtccgtgtactcccaagaatctacgcacgcgagacagga	3241
Db	3238	ATTCGAAACCTTATAGCAGGGCGCTCCGCTTGACCCATCTTCCTGGCCCGGACCGCGA	3297
OY	3242	tgaaccgagacagagaaacatcgcggttgcgttgtaagcaatccctggacccacagtaacga	3301
Db	3298	CGAACCCGGAACACAGACACTCCAGGGCTGGCTTACGAGATCTCGAACTC--AGCACGA	3355
OY	3302	ccgagacatcgaggtgtccggaatctctcacaatgactcgaattgaaacaaacactgggga	3361
Db	3355	GGGGGACATGGAAGTGGCGGAGCTGTTGGATGATACAGATCGGGAAGTCACGGGAGA	3414
OY	3362	acttgaagacgcgatgtgacttgaagggggtattggggacattacgaatatctatttgagc	3422
Db	3415	GCTGGAAACGCCATGTGACCTCAAGGAGATATGGGGAGTACGTACGTATACATATACGGGC	3474
OY	3422	cttcgacacagcgctacccgcagacagagttccaacagagctacacccctgtcatccgcc	3481
Db	3475	ATTGACACCGAGGCAATTCGGCAATGTGTCATGAACGAGACATATGAGTGTGATCATCTAC	3534
OY	3482	ctacacactccacacccctgtgtgtctgtctccgcaaacccgactccgtcatctcgcttc	3541
Db	3535	GTTCACTACTAGCGCGCTGATGGTGTCTTCCCGACCAACGATGCGCTCATACGACTTGC	3594
OY	3542	taggagcgcgcgaacagaagcgcgtctctgcygagcgtgcgaacagttctcttgaagc	3601
Db	3595	GAGGGAACGAGCTGTATTCATGATGAGTGTTCAGGACAGTGAACGAGATTTCTTGGACGG	3655
OY	3602	gattctacgcacacgaaggaagcgcctcgaacgcgggcagcgtcgaacttccacgtccaggg	3661
Db	3655	GATATCGCGCGACTGATCCGAGCGAGACTCCACGCGGGCGTGTCTACCTTCCAAAGTGTAGG	3714
OY	3662	aaatgaagagagcgcttgcagactttgataataacttgaaagtggagacagagaaatatagcg	3721
Db	3715	CGATGAGGATCACACAGGTACTTTCAAGTAGTTAAAGATGEC-----GAACACTGG	3766
OY	3722	gcagctatatactaacgcagttcttccagagcaatcagacaatttcagatcagatccg	3781
Db	3769	CTCGTTAGGTTACTGCAACCGTTCCAGAGGAGATACGAGGATTCGGATTAACATTCG	3828
OY	3782	ggcacaagagcgcgcgcaggaacccggtcccgctttgagaccgaagttcagtttgcgttgt	3844
Db	3829	CGCTACAGACCAAGGAACCGACCGGACCGCTGTCCACGAGATGACCTTCAAGATGT	3886
OY	3842	cttcgtaccacaagaagcgagaccagtggtctcagcgaataatgcagctactgtccctct	3901
Db	3889	TTTTGTGCCACGCAAGGAGAACCTAGATTCCGGGTCTCTGAAACATGCTGTGGCTTTCA	3944
OY	3902	cgaaggtgaagaagcctcgtgtgagagtttgagctgcgcgaagcagaagaccttaaaa	3961
Db	3949	AGAAAAGAGTCCCGCATGTGAAAGTCTCACCAACTCTCTTAGCACAGACATCAAGAA	4006
OY	3962	ccacctctgcgaataatgactgcgaagataatcactacacagttatttgaagcgaacaga	4021
Db	4009	CCATCTCTGTGAACGACATGCTCACACCAATTTACTATCTGATTAATGATGGCAACGCA	4066
OY	4022	gggtcttctgtacttgacacagttcaagcagcgtcatctccctctgcgcagagttgagccg	4081

[illegible]

[illegible]

US-08-272-882D-1

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 15, 2002, 16:00:42 ; Search time 61.22 Seconds
(Without alignments)
631.136 Million cell updates/sec

Title: US-09-715-909-2

Perfect score: 8909

Sequence: 1 MGVERFPVAVLVASALASAL.....THGNFGRKSTPYLPQPHK 1717

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

212252 seqs, 22503292 residues

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents -AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4693	52.7	1528	1	US-08-326-117B-2
2	4693	52.7	1528	3	US-08-982-129-2
3	459	5.2	119	1	US-08-326-117B-10
4	459	5.2	119	3	US-08-982-129-10
5	457.5	5.1	138	1	US-08-326-117B-12
6	457.5	5.1	138	3	US-08-982-129-12
7	398	4.5	146	1	US-08-326-117B-15
8	398	4.5	146	3	US-08-982-129-15
9	382.5	4.3	122	1	US-08-326-117B-14
10	382.5	4.3	122	3	US-08-982-129-14
11	376	4.2	120	1	US-08-326-117B-16
12	376	4.2	120	3	US-08-982-129-16
13	363.5	4.1	113	1	US-08-326-117B-8
14	363.5	4.1	113	3	US-08-982-129-8
15	352	4.0	124	1	US-08-326-117B-13
16	352	4.0	124	3	US-08-982-129-13
17	337.5	3.8	112	1	US-08-326-117B-17
18	337.5	3.8	112	3	US-08-982-129-17
19	324	3.6	1026	1	US-07-998-003A-95
20	324	3.6	1026	1	US-08-453-274B-95
21	324	3.6	1026	1	US-08-453-695A-95
22	324	3.6	1026	1	US-08-268-161A-95
23	324	3.6	1026	2	US-08-453-702A-95
24	324	3.6	1026	4	US-09-099-639-95
25	324	3.6	1026	5	PCT-US93-12588-95
26	324	3.6	1026	5	PCT-US95-08071-95
27	324	3.6	1203	1	US-07-998-003A-103

28	324	3.6	1203	1	US-08-453-274B-103	Sequence 103, App
29	324	3.6	1203	1	US-08-453-695A-103	Sequence 103, App
30	324	3.6	1203	1	US-08-268-161A-103	Sequence 103, App
31	324	3.6	1203	2	US-08-453-702A-103	Sequence 103, App
32	324	3.6	1203	4	US-09-099-639-103	Sequence 103, App
33	324	3.6	1203	5	PCT-US93-12588-103	Sequence 103, App
34	324	3.6	1203	5	PCT-US95-08071-103	Sequence 103, App
35	301.5	3.4	106	1	US-08-326-117B-9	Sequence 9, App1
36	301.5	3.4	106	3	US-08-982-129-9	Sequence 9, App1
37	298	3.3	104	1	US-08-326-117B-11	Sequence 11, App1
38	298	3.3	104	3	US-08-982-129-11	Sequence 11, App1
39	297.5	3.3	979	4	US-08-878-474-5	Sequence 5, App1
40	295.5	3.3	797	1	US-08-453-695A-112	Sequence 112, App
41	295.5	3.3	797	1	US-08-268-161A-112	Sequence 112, App
42	295.5	3.3	797	2	US-08-453-702A-112	Sequence 112, App
43	295.5	3.3	797	4	US-09-099-639-112	Sequence 112, App
44	295.5	3.3	797	5	PCT-US95-08071-112	Sequence 112, App
45	280.5	3.1	555	2	US-08-453-702A-98	Sequence 98, App1

ALIGNMENTS

RESULT 1
US-08-326-117B-2
Sequence 2, Application US/08326117B
Patent No. 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
INVENTOR: JET TAB
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-326-117B-2

Query Match 52.7%; Score 4693; DB 1; Length 1528;
Best Local Similarity 60.7%; Pred. No. 0;
Matches 925; Conservative 220; Mismatches 357; Indels 22; Gaps 13;

1 MGVERFPVAVLVASALANORCSYIAIPR-PEMPELPPIIDYEGKSNSEOPPLIGPT 59
1 MAVVRIAPFLVIAIAVLAQKCGTWTATIPRLPRDNLFPVLFEGQVMSORPLPAPE 60

QY 60 REEVOEMNF---LPDOMIOVIMEEIEGDVIAKLANTOGSNTPLVSI--MSGOPRAOLG 114
D 61 RDDCLMDAYHITANLGOVYIYMEDEIETITALLNNGSTPIELPFLSGSNTL-LM 119
QY 115 PEFRONEADGWSLVITQRODYETATMQSYVESIOVEGSOAVLVALEIVINDNPILQ 174
D 120 PVIRR--VDNGSASHHNRHGYELPGMOQYMFNVAVDQSLVAGVSLAIVINDNAPILQ 177
QY 175 VVSACVPEHEEARLITDCVOYVSDRDEISTREPMFRVDSRADDESIFVANGVEIDPSDW 234
D 178 NPEORCVPELDEPGLTECYOVSDADGRISTEFMFTRIDSVR-GDEEFYIETERNIPQW 236
QY 235 FNMKTGINSPLNEETQOLHFSVTASDSLPNNHTVTMMOVENESRPRMVFESVQ 294
D 237 MMLNMTIIVNTSLNFVISPFIHFSVTALDSLPNHTVTMMOVAVNNSRPRMVEIFAVQ 296
QY 295 QFDEKTNOSFLRAIDGDTGINRAINYTLIRDDADDFESLEVI--EDGAILVHTEIDRD 351
D 297 OFEESYQNFVRAIDGDTINMPINRYLTINEEDTFESIEALPGKSGAVFLVSPIDRD 356
QY 352 KLERELMTITVAKSPTDASFATEHIFIIYDNDVQDRPEPLHKEYSIDIMEETMTNF 411
D 357 TLQREVEPLTIVAKYDEAFSTSNVAVITVDINDQRPEDPHKEYRLAIMEEPLTNF 416
QY 412 NEEFGRHDDRGEMAYVELEDFVPPGASAFYIAPGSGYQROTFFINGTINHTMLDYED 471
D 417 DKERGFHKDLCQANQYVRLSVYDPCGAEEFYIAPVGVQROTFINGTINHSMLDYED 476
QY 472 VIFONIIKAVADNNAHSHVGEALVYVNLINMDELPIFEESYSASFSETVGAGFPVA 531
D 477 PEFOSTIRVAVATDNDNRHGVAVLHIDLINMDEQPIFEHAAVQTVFDETEGGEFFVA 536
QY 532 TVLADRIDDVVYVSLMGNAVDTLFIDESGETFVSMDDAFDYROMTLFOVADPFL 591
D 537 KAVADRIDGVEVETLLGNAVNFLLTDLKLGDIKVSANDSNHRESELEFQVAVATPL 596
QY 592 GDGPHNTVTTQVLEIEEDVNTNPTLRPLRSTPSVEENVPEGEISREITATDPPTSAYL 651
D 597 GE-PHHTQSOLVIRLNDINNTPTLRPLRPSQVEENVPGCHVITQGLRATDPTTADL 655
QY 652 WPEIDWSTWATKOGRETNPEYVGCIVETIYPTEGNRGSAIGRLVQVEIRDVNTDOFE 711
D 656 RFEIIMWDSFATKQGRANDEFRCVCIETIFPEINNRKGLAIGHVAREIRHNVTIDYE 715
QY 712 EEFEMXLTVARVADNTVYGDYDEATFTITIDMNDNAPFANGTLQTMVRELAASGT 771
D 716 EFEVSLTVARVADNTVYGDYDESMLTITITIDMNDNAPVWVEGLEONFRERESAGGL 775
QY 772 LIGSVLATIDIGPLXNQVRYTIOPRNNTPREGIYKIDFTTGOIEVDANEAIDADEPFRYL 831
D 776 VVGSVRADIDIGPLXNQVRYTIOPRNNTPREGIYKIDFTTGOIEVDANEAIDADEPFRYL 835
QY 832 YTYVYASDECSLENTECPDPSNTEFEVGDIEIETIDNNKVPREPLTEKFNVTYVWENA 891
D 836 YTYVYASDRCTEPPADCPDPTVWETEGNTITHTDNNKVPQAEETKFKFTVYVYENA 895
QY 892 TSGDEVOLYSHDRDRDELTVRYRTNMFVAVNPRLDFFEVDLQGRLEVHP--GDEK 948
D 896 THLDEVVLLIASDLDRDITHTVSYVINYAVNPRLMNEFSVNRKGLVYVYDETOGSGSEV 955
QY 949 LDRDDEPTHTIEVNFIDFESDGGRRNODEVELEFVLLVDVNDNAPMPPLDELRFDVS 1008
D 956 LDRDDEPTHTIEVNFIDFESDGGRRNODEVELEFVLLVDVNDNAPMPPLDELRFDVS 1015
QY 1009 EGAVAGVAVLEIYAPDDEPDITNSRVYGLIDLTITDROIEVDOLFTMISIEKKTGEL 1068
D 1016 ENLKGVARLEPHIFAPDDEPDITNSRVYGLIDLTITDROIEVDOLFTMISIEKKTGEL 1074
QY 1069 ETAMDLRGYWGVEFFIAPFHGYPQORSNETYTLVIRPNVHHVFPQDVSITRISR 1128
D 1075 ETAMDLRGYWGVEFFIAPFHGYPQORSNETYTLVIRPNVHHVFPQDVSITRISR 1134

QY 1129 ERATEGVLAATANEFLERIPVATDEDDGLHAGSVTFHVQNEEAVQYFDITEVGAGENSQ 1168
D 1135 ERAVINGVLAVNGEFLERISATIDPDGLHAGVTFEQVYVDESSQRYEQV--VNDGENLS 1192
QY 1189 LILROLPEQIROFRITIRATDGTCEPGLMTDVTFSVFVPTQGDVPVSENAATVAFE 1248
D 1193 LRLQAVPEERIEFRITIRATDGTCEPGLMTDVTFSVFVPTQGDVPVSENAATVAFE 1252
QY 1249 GEEGRSEFELPOAMDELKCNHCEDDCCODIYREFIDGNNEGFLVLDQSSNVLSLOELDR 1308
D 1253 KSAGNEESHQPLADODINNHCEDDCHSITRILIDGNSGEGFLDPPVNRRLFLKELIRE 1312
QY 1309 VATSYTLHIAASNSPDATGIPLOTISILVTVVNVREANRPFEODLTAGISTIDSTIRE 1368
D 1313 QSASHITQVLAASNSPDG-GIPLPASILTVTVYVREADRPVFRVRELYTAGISTIDSTIRE 1371
QY 1369 LITVRASTEDDTTYTTRDRASMOIDSSLEAVRDSAFALHATTGYLSLNMPTASMGME 1428
D 1372 LRLHATQSESAITVYADYDTMVVDPSELAVROSAPVLAQOTGLVTLINPTATMGLF 1431
QY 1429 EEDVATATASADIDTARVYVYLISQNRVTFIFDQLETVEQNRFFIATSTGFNMTCN 1468
D 1432 KFEVATATAGAQDRTDVTYVYVSSQNRVYFVNTQOVEDNRDIDTISAGFNMTCN 1491
QY 1489 IDQVVPFSDS--SGVAQDDTTEVRA 1511
D 1492 IDQVVPANDPVTGVALHSTQMAA 1515

RESULT 2
US-08-982-129-2
: Sequence 2, Application US/08982129
: Patent No. 6007981
: GENERAL INFORMATION:
: APPLICANT: BULLA, LEE A.
: TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: MORRISON & FOERSTER
: STREET: 2000 Pennsylvania Ave. N.W.
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20006-1812
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/982,129
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/326,117
: FILING DATE: 19-OCT-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: MILLMAN, ROBERT A
: REGISTRATION NUMBER: 36,217
: REFERENCE/DOCKET NUMBER: 7112-0037.00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 887-1500
: TELEFAX: (202) 887-0763
: TELEX: 90-4030
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1528 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-982-129-2

Query Match 52.7%: Score 4693: DB 3: Length 1528;
 Best Local Similarity 60.7%: Pred. No. 0:
 Matches 925; Conservative 220; Mismatches 357; Indels 22; Gaps 13:

```

QY 1 MGEVPEPVLVSLASALANORCSYIIAIPR-PETPELPIDYEGKMSSEQLIPGPT 59
DB 1 MAVDVIAFLVFLVIAPAVLAORCGYMTAIPRLPRDNLPLYNFGQWMSQRLPAPR 60
QY 60 REVCENENF---LPDMQIYIIMEEIEGDVITAKUNQGSNTPVUST--MSGPRAQIG 114
DB 61 RDLCDKADAHVITANLGTQIYIMDEIEDEITAILNYPNSPTPELPELSSSYML-LM 119
QY 115 PERROMADGQMSLVITODDYETATMGOSVPSIOVEGSGQALVALLEYINDNDNPILQ 174
DB 120 PYIAR--VUNGASHHAKHKELEPMQOQMEVNRVDGSLVAGSLAVINDDNAPITQ 177
QY 175 VVSACVYPERGEARLDCYOVSDRDEISTREMPFVDSRRADESIFVMGEYPSDM 234
DB 178 NPEPCVAPPELGEPLGECYQVSDADGRISTEMTRIDSVR-GDEETFYIERTNIPNQM 236
QY 235 PNMKMTVGINSPLNETTQULHIFSVTASDLPNNHTVMQVQVENESPPRWELFSVQ 294
DB 237 MMLNMTIGVNTSLNFTSPLHIFSVTALDLSLPTHTVMQVQVANSRPRLLEIFAVO 296
QY 295 OPEDEKNOSESLAIDGDCINRAINYTLIRDDADDFSSLEYI---EDGAILHYTIDND 351
DB 297 OPEEKQNTVYVAIDGDEIINMPINRYLITNEDEFFSIEALPGKSGAVELVSPIDND 356
QY 352 KLERELNLTIVAKSSTDAFAEHIFFITVNDVNDQPREPLHRESIDIMEETPTLNF 411
DB 357 TLQREVPILYAKKDEAFSTSTNVIIIVNDIQREPRHKEVRLAMETPTLINF 416
QY 412 NEEFGHDLGEBNAQYVELEDPERPGAASAFYIAPSGGYORQFIMGTINHMLDYED 471
DB 417 DKEFGHDLGCONAQYVLESVPDGAFAFYIAPSGGYORQFIMGTINHMLDYED 476
QY 472 VLFONTIIVKAVDMMNASHGEALVYVNLIMNDELFFEESSYSASKEVYVAGAFPA 531
DB 477 PEFQSTITIVATDNDTGHVAGALHIDLIMNDEQPLFEHAVQVITDETEGEEFFVA 536
QY 532 TVALDRDIDDDVYVSHLMANAVYLEFIDESTGEIEFVSMADAFYHROMLTFQVADRDTL 591
DB 537 KAVAHROIDGVYEHLLGNAYNELIDKLTDGIRKSANDSNYHRESLEFQVADRDTL 596
QY 592 GGGPHNTVTLQVIELEDVNNPTPLRLRSPSVSEENVPEGEYSIRELTATDPDTSAVL 651
DB 597 GE-PFHATASQIVIRLNDINNTPLRLRPGSPQVEENVPDGHVITQELRATDDPTADL 655
QY 652 WFEIDMDSTWATKQGEETPTEYVGCIVLETIYPTBGNAGSAIGRLVQVEIDNVTIDPE 711
DB 656 REEIMNDTSFATKQGAOPDEFRCVEIETIFPELINNGLAIGRVAREIRHNVTIDPE 715
QY 712 EPEMLYLTFRVBDLNTVIDDDYDEAFETTTIIDMNDNAPFANGTLTOMARYELASGT 771
DB 716 EEEVLSLTYRVADLNTYVDDYDESMITTTIIDMNDNAPVWEGTLEQVFRRESAGSL 775
QY 772 LIGSVLANDIDGPLYQVRYTIQPRNNTPEGLVKIDFTTGQIEVDANEADADEPMREYL 831
DB 776 VVGSVYRADIDGPLYQVRYTIQPRNNTPEGLVKIDFTTGQIEVDANEADADEPMREYL 835
QY 832 YXTVYIASDSCSLENTECPDSNTEFEVDDIEIETIDNNKVPPEPLTEKFNFTTYVWENA 891
DB 836 YXTVYIASDSCSLENTECPDSNTEFEVDDIEIETIDNNKVPPEPLTEKFNFTTYVWENA 895
QY 892 TSGDEVOLYISDRDDELHYHTFRYTMNPAVNRPLDFEVDLDGRLEFVHP---GDEK 948
DB 896 THLDEVYTLIADDLDRDELHYHTFRYTMNPAVNRPLDFEVDLDGRLEFVHP---GDEK 955
QY 949 LDRDGEPTHTTFVNEIDNFSQDGRNQDEVEIFVLLVDNANAPENPLRDELRFVDS 1008
DB 956 LDRDGEPTHTTFVNEIDNFSQDGRNQDEVEIFVLLVDNANAPENPLRDELRFVDS 1015

```

```

QY 1009 EGAVAGRVLPETIYAPDRDEPPDNRSGVGIILDTITRDIEVPLIFTMISLENKTEL 1068
DB 1016 ENIKQCVRLPEPHIFAPDRDEPDIDNSRVGYELIUNLS-TERDIEVELEPMIOTANTGEL 1074
QY 1069 ETAMDLRGWGTYELEFEAFDGHGPOORSNETYTLVIRPNFHHFVFPPOPDSYIRLSR 1128
DB 1075 ETAMDLRGWGTYAIHIRAFDGHGIPQMSNETYELIHPFNVAPEFVPTNDAYIRLAR 1134
QY 1129 ERATEGVATANELEFIYATDEDEGLHAGSVYTRHVGSENEAVQYFDTTEVGAGENSQ 1188
DB 1135 ERAYINGVATANGELERISATIDPDGLHAGVYTFQVWDEDSORFYOV--VNDGENLS 1192
QY 1189 LILRQLEPQIOHOFRTTITATGTEPEPLMTDVFVSVYFVTCQGPVSENAATYAFPE 1248
DB 1193 LRLQAVPEIEKEFRITATDQGDIDPGLSIDMTFRVYFVTCQGPVSENAATYAFPE 1252
QY 1249 GEGGLNESEFELQAEDELKKNLCEDDQDIYRFRIDGNBGLFVLDOSSNVIslaELDRE 1308
DB 1253 KASGMESHQPLAQDIKNHLCEDDCHSYRRIIDGNSGHRGLDPVRMRFLKKELINE 1312
QY 1309 VATSTYLIHAASNPDAIGIPLOTSLVYTVANREANRPITEBOLYTAGISTLSIGRE 1368
DB 1313 QASHTLQVAASNSPDG-GIPLASITVTVVREADPRPVFVRELITAGISTLSIGRE 1371
QY 1369 LILVASHTEDDITTYTIDRASMOGLSLEAVRDSAFALHATGVLSLMOPTASHMGF 1428
DB 1372 LILVASHTEDDITTYTIDRASMOGLSLEAVRDSAFALHATGVLSLMOPTASHMGF 1431
QY 1429 EFDVATIPASAIIDTARVAVYLISSONRYTFLPDQLETFEONRNFIAFTSTGNMTCN 1488
DB 1432 KEVYATVATAGADRDYVYVSSONRYTFLPDQLETFEONRNFIAFTSTGNMTCN 1491
QY 1489 IDQVVPFSDS-SGVAODTTEYVA 1511
DB 1492 IDQVVPANDPVYVVALEHSTQMAA 1515

```

RESULT 3
 US-08-326-117B-10
 ; Sequence 10, Application US/08326117B
 ; Patent No. 5693491
 ; GENERAL INFORMATION:
 ; APPLICANT: BULLA, LEE A.
 ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/326,117B
 ; FILING DATE: 19-OCT-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MILLMAN, ROBERT A.
 ; REGISTRATION NUMBER: 36,217
 ; REFERENCE/DOCKET NUMBER: 7112-0037.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-326-117B-10

Query Match 5.2%; Score 459; DB 1; Length 119;
Best Local Similarity 71.4%; Pred. No. 1e-31;
Matches 85; Conservative 15; Mismatches 19; Indels 0; Gaps 0;

QY 393 LAKEYSIDIMETPTLNFNEGFHDLGQNAQYVLESDVPPGAASAFYIAPGSGY 452
DB 1 IHKEYRLAIMEETPLTNFDEKFGHDKDLGQNAQYVRLSDVPPGAASAFYIAPGSGY 60
QY 453 ORQPTFMGTINHTMDYEDVIFQNIIRKAVDMNASHGVALYVVLINMNDLPPIF 511
DB 61 ORQPTFMGTINHTMDYEDVIFQNIIRKAVDMNASHGVALYVVLINMNDLPPIF 119

US-08-982-129-10

Sequence 10, Application US/08982129

Patent No. 6007981

GENERAL INFORMATION:

APPLICANT: BULLA, LEE A.

APPLICANT: JI, TAE

TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS

TITLE OF INVENTION: TOXIN

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/982,129

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/326,117

FILING DATE: 19-OCT-1994

TORNEY/AGENT INFORMATION:

NAME: MILLMAN, ROBERT A

REGISTRATION NUMBER: 36,217

REFERENCE/DOCKET NUMBER: 7112-0037.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-982-129-10

Query Match 5.2%; Score 459; DB 3; Length 119;

Best Local Similarity 71.4%; Pred. No. 1e-31;

Matches 85; Conservative 15; Mismatches 19; Indels 0; Gaps 0;

QY 393 LAKEYSIDIMETPTLNFNEGFHDLGQNAQYVLESDVPPGAASAFYIAPGSGY 452
DB 1 IHKEYRLAIMEETPLTNFDEKFGHDKDLGQNAQYVRLSDVPPGAASAFYIAPGSGY 60
QY 453 ORQPTFMGTINHTMDYEDVIFQNIIRKAVDMNASHGVALYVVLINMNDLPPIF 511
DB 61 ORQPTFMGTINHTMDYEDVIFQNIIRKAVDMNASHGVALYVVLINMNDLPPIF 119

DB 1 IHKEYRLAIMEETPLTNFDEKFGHDKDLGQNAQYVRLSDVPPGAASAFYIAPGSGY 60
QY 453 ORQPTFMGTINHTMDYEDVIFQNIIRKAVDMNASHGVALYVVLINMNDLPPIF 511
DB 61 ORQPTFMGTINHTMDYEDVIFQNIIRKAVDMNASHGVALYVVLINMNDLPPIF 119

RESULT 5

US-08-326-117B-12

Sequence 12, Application US/08326117B

Patent No. 5693491

GENERAL INFORMATION:

APPLICANT: BULLA, LEE A.

APPLICANT: JI, TAE

TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS

TITLE OF INVENTION: TOXIN

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/326,117B

FILING DATE: 19-OCT-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MILLMAN, ROBERT A

REGISTRATION NUMBER: 36,217

REFERENCE/DOCKET NUMBER: 7112-0037.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 138 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-326-117B-12

Query Match 5.1%; Score 457.5; DB 1; Length 138;
Best Local Similarity 64.7%; Pred. No. 1.8e-31;
Matches 90; Conservative 22; Mismatches 22; Indels 5; Gaps 2;

QY 618 RLPRSTPSVENVPEGEISREITATDPDTSAYLWPEIDWSTWATKQGRNTPTEYGC 677
DB 1 RLPRSPQVEERVPAHAYTQELRATDPDTADLRFETINWDSFATKQGRANPDPEFRC 60
QY 678 IYIETIYPTL-----GNRSATGRLVVOEIRDVNTIDEEFEMLYTVVRDINTYIGDY 733
DB 61 VEIETIYPTL-----GNRSATGRLVVOEIRDVNTIDEEFEMLYTVVRDINTYIGDY 119
QY 734 DEATFTITIDMNDNAPIF 752
DB 120 DESMLTITIDMNDNAPYV 138

RESULT 6

US-08-982-129-12

Sequence 12, Application US/08982129

Patent No. 6007981

GENERAL INFORMATION:

APPLICANT: BULLA, LEE A.


```

ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-982-129-15

Query Match 4.5%; Score 398; DB 3; Length 146;
Best Local Similarity 53.4%; Pred. No. 2,4e-26;
Matches 78; Conservative 14; Mismatches 26; Indels 28; Gaps 1.

998 PLEPDLRFDVSEGAVAGRYVLPETIAPRDRDEPTDNRVGYGIDLT----- 1045
1 PPSPELSTWISSENKOGVRLPHEPIFAPRDEPTDNRVGYEILNSTERDIEVPELFW
DB 1046 -----PDRDIEVPLFTMISTENKTGELETAMDLRGNGYEIFEARD 1089
Db 61 IQIIANVTGYEILNSTERDIEVPELFWIQIANVTGELETAMDLRGNGYTAIYLARD 120
OY 1090 HGYPQORSNETYTVIRPYNFHHPVF 1115
Db 121 HGIPQMSMRETYELIHPNYAPPEF 146

RESULT 9
US-08-326-117B-14
Sequence 14 Application US/08326117B
Patent No. 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids
TYPE: amino acid
STRANDEDNESS: single

```

```

      TOPOLOGY: linear
      US-08-326-117B-14

      Query Match
      Best Local Similarity 59.5%; Pred. No. 3.7e-25;
      Matches 72; Conservative 18; Mismatches 30; Indels 1; Gaps 1;

      QY 878 TERKNTVYVWENATSGDEVVQLYSHDRDRDELHYTRYTNFANVNPRLRDFEVDLTG 937
      Db 2 TTKKDFVYIYENATHLDEVTLLASDLDRDEIYHMSYINAVNPRLMFESVNETG 61
      QY 938 RLEHYHPGD-EKLDRDGEPTHTIFVNEIDNFSDGGRNQDEVEIFVLLDVNDNAPE 996
      Db 62 LVYDYERQSGSLDRDGEPTHTRIFFENLIDNFMGEGCGRNQNDTEVLVILLDVNDNAPE 121
      QY 997 M 997
      Db 122 L 122

      RESULT 10
      US-08-982-129-14
      ; Sequence 14, Application US/08982129
      ; Patent No. 6007981
      ; GENERAL INFORMATION:
      ; APPLICANT: BULLA, LEE A.
      ; APPLICANT: JT, TAE
      ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
      ; TITLE OF INVENTION: TOXIN
      ; NUMBER OF SEQUENCES: 26
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: MORRISON & FOERSTER
      ; STREET: 2000 Pennsylvania Ave. N.W.
      ; CITY: Washington
      ; STATE: DC
      ; COUNTRY: USA
      ; ZIP: 20006-1812
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: Patent In Release #1.0, Version #1.30
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/982,129
      ; FILING DATE:
      ; CLASSIFICATION:
      ; PRIOR APPLICATION DATA:
      ; APPLICATION NUMBER: US 08/326,117
      ; FILING DATE: 19-OCT-1994
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: MILLMAN, ROBERT A.
      ; REGISTRATION NUMBER: 36,217
      ; REFERENCE/DOCKET NUMBER: 7112-0037.00
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: (202) 887-1500
      ; TELEFAX: (202) 887-0763
      ; TELETYPE: 90-4030
      ; INFORMATION FOR SEQ ID NO: 14:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 122 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS: single
      ; TOPOLOGY: linear
      ; US-08-982-129-14

      Query Match
      Best Local Similarity 59.5%; Pred. No. 3.7e-25;
      Matches 72; Conservative 18; Mismatches 30; Indels 1; Gaps 1;

      QY 878 TERKNTVYVWENATSGDEVVQLYSHDRDRDELHYTRYTNFANVNPRLRDFEVDLTG 937
      Db 2 TTKKDFVYIYENATHLDEVTLLASDLDRDEIYHMSYINAVNPRLMFESVNETG 61
      QY 938 RLEHYHPGD-EKLDRDGEPTHTIFVNEIDNFSDGGRNQDEVEIFVLLDVNDNAPE 996
      Db 62 LVYDYERQSGSLDRDGEPTHTRIFFENLIDNFMGEGCGRNQNDTEVLVILLDVNDNAPE 121
      QY 997 M 997
      Db 122 L 122

      RESULT 10
      US-08-982-129-14
      ; Sequence 14, Application US/08982129
      ; Patent No. 6007981
      ; GENERAL INFORMATION:
      ; APPLICANT: BULLA, LEE A.
      ; APPLICANT: JT, TAE
      ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
      ; TITLE OF INVENTION: TOXIN
      ; NUMBER OF SEQUENCES: 26
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: MORRISON & FOERSTER
      ; STREET: 2000 Pennsylvania Ave. N.W.
      ; CITY: Washington
      ; STATE: DC
      ; COUNTRY: USA
      ; ZIP: 20006-1812
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: Patent In Release #1.0, Version #1.30
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/982,129
      ; FILING DATE:
      ; CLASSIFICATION:
      ; PRIOR APPLICATION DATA:
      ; APPLICATION NUMBER: US 08/326,117
      ; FILING DATE: 19-OCT-1994
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: MILLMAN, ROBERT A.
      ; REGISTRATION NUMBER: 36,217
      ; REFERENCE/DOCKET NUMBER: 7112-0037.00
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: (202) 887-1500
      ; TELEFAX: (202) 887-0763
      ; TELETYPE: 90-4030
      ; INFORMATION FOR SEQ ID NO: 14:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 122 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS: single
      ; TOPOLOGY: linear
      ; US-08-982-129-14

      Query Match
      Best Local Similarity 59.5%; Pred. No. 3.7e-25;
      Matches 72; Conservative 18; Mismatches 30; Indels 1; Gaps 1;

      QY 878 TERKNTVYVWENATSGDEVVQLYSHDRDRDELHYTRYTNFANVNPRLRDFEVDLTG 937
      Db 2 TTKKDFVYIYENATHLDEVTLLASDLDRDEIYHMSYINAVNPRLMFESVNETG 61
      QY 938 RLEHYHPGD-EKLDRDGEPTHTIFVNEIDNFSDGGRNQDEVEIFVLLDVNDNAPE 996
      Db 62 LVYDYERQSGSLDRDGEPTHTRIFFENLIDNFMGEGCGRNQNDTEVLVILLDVNDNAPE 121
      QY 997 M 997
      Db 122 L 122

      RESULT 10
      US-08-982-129-14
      ; Sequence 14, Application US/08982129
      ; Patent No. 6007981
      ; GENERAL INFORMATION:
      ; APPLICANT: BULLA, LEE A.
      ; APPLICANT: JT, TAE
      ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
      ; TITLE OF INVENTION: TOXIN
      ; NUMBER OF SEQUENCES: 26
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: MORRISON & FOERSTER
      ; STREET: 2000 Pennsylvania Ave. N.W.
      ; CITY: Washington
      ; STATE: DC
      ; COUNTRY: USA
      ; ZIP: 20006-1812
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: Patent In Release #1.0, Version #1.30
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/982,129
      ; FILING DATE:
      ; CLASSIFICATION:
      ; PRIOR APPLICATION DATA:
      ; APPLICATION NUMBER: US 08/326,117
      ; FILING DATE: 19-OCT-1994
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: MILLMAN, ROBERT A.
      ; REGISTRATION NUMBER: 36,217
      ; REFERENCE/DOCKET NUMBER: 7112-0037.00
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: (202) 887-1500
      ; TELEFAX: (202) 887-0763
      ; TELETYPE: 90-4030
      ; INFORMATION FOR SEQ ID NO: 14:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 122 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS: single
      ; TOPOLOGY: linear
      ; US-08-982-129-14

      Query Match
      Best Local Similarity 59.5%; Pred. No. 3.7e-25;
      Matches 72; Conservative 18; Mismatches 30; Indels 1; Gaps 1;

      QY 878 TERKNTVYVWENATSGDEVVQLYSHDRDRDELHYTRYTNFANVNPRLRDFEVDLTG 937
      Db 2 TTKKDFVYIYENATHLDEVTLLASDLDRDEIYHMSYINAVNPRLMFESVNETG 61
      QY 938 RLEHYHPGD-EKLDRDGEPTHTIFVNEIDNFSDGGRNQDEVEIFVLLDVNDNAPE 996
      Db 62 LVYDYERQSGSLDRDGEPTHTRIFFENLIDNFMGEGCGRNQNDTEVLVILLDVNDNAPE 121
      QY 997 M 997
      Db 122 L 122

      RESULT 10
      US-08-982-129-14
      ; Sequence 14, Application US/08982129
      ; Patent No. 6007981
      ; GENERAL INFORMATION:
      ; APPLICANT: BULLA, LEE A.
      ; APPLICANT: JT, TAE
      ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
      ; TITLE OF INVENTION: TOXIN
      ; NUMBER OF SEQUENCES: 26
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: MORRISON & FOERSTER
      ; STREET: 2000 Pennsylvania Ave. N.W.
      ; CITY: Washington
      ; STATE: DC
      ; COUNTRY: USA
      ; ZIP: 20006-1812
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: Patent In Release #1.0, Version #1.30
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/982,129
      ; FILING DATE:
      ; CLASSIFICATION:
      ; PRIOR APPLICATION DATA:
      ; APPLICATION NUMBER: US 08/326,117
      ; FILING DATE: 19-OCT-1994
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: MILLMAN, ROBERT A.
      ; REGISTRATION NUMBER: 36,217
      ; REFERENCE/DOCKET NUMBER: 7112-0037.00
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: (202) 887-1500
      ; TELEFAX: (202) 887-0763
      ; TELETYPE: 90-4030
      ; INFORMATION FOR SEQ ID NO: 14:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 122 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS: single
      ; TOPOLOGY: linear
      ; US-08-982-129-14

      Query Match
      Best Local Similarity 59.5%; Pred. No. 3.7e-25;
      Matches 72; Conservative 18; Mismatches 30; Indels 1; Gaps 1;

      QY 878 TERKNTVYVWENATSGDEVVQLYSHDRDRDELHYTRYTNFANVNPRLRDFEVDLTG 937
      Db 2 TTKKDFVYIYENATHLDEVTLLASDLDRDEIYHMSYINAVNPRLMFESVNETG 61
      QY 938 RLEHYHPGD-EKLDRDGEPTHTIFVNEIDNFSDGGRNQDEVEIFVLLDVNDNAPE 996
      Db 62 LVYDYERQSGSLDRDGEPTHTRIFFENLIDNFMGEGCGRNQNDTEVLVILLDVNDNAPE 121
      QY 997 M 997
      Db 122 L 122

      RESULT 10
      US-08-982-129-14
      ; Sequence 14, Application US/08982129
      ; Patent No. 6007981
      ; GENERAL INFORMATION:
      ; APPLICANT: BULLA, LEE A.
      ; APPLICANT: JT, TAE
      ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
      ; TITLE OF INVENTION: TOXIN
      ; NUMBER OF SEQUENCES: 26
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: MORRISON & FOERSTER
      ; STREET: 2000 Pennsylvania Ave. N.W.
      ; CITY: Washington
      ; STATE: DC
      ; COUNTRY: USA
      ; ZIP: 20006-1812
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: Patent In Release #1.0, Version #1.30
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/982,129
      ; FILING DATE:
      ; CLASSIFICATION:
      ; PRIOR APPLICATION DATA:
      ; APPLICATION NUMBER: US 08/326,117
      ; FILING DATE: 19-OCT-1994
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: MILLMAN, ROBERT A.
      ; REGISTRATION NUMBER: 36,217
      ; REFERENCE/DOCKET NUMBER: 7112-0037.00
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: (202) 887-1500
      ; TELEFAX: (202) 887-0763
      ; TELETYPE: 90-4030
      ; INFORMATION FOR SEQ ID NO: 14:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 122 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS: single
      ; TOPOLOGY: linear
      ; US-08-982-129-14

      Query Match
      Best Local Similarity 59.5%; Pred. No. 3.7e-25;
      Matches 72; Conservative 18; Mismatches 30; Indels 1; Gaps 1;

      QY 878 TERKNT
```

Db 2 TTKFDTVYIYNATHLDEVVTLIASDLDRBEIYHMSVYIYAVNPLMFESVNRFTG 61
QY 938 RLEVIYIPBD-EKLDGDGDEPHITIVNFIDNFFSDGGRNODVEIYVLLDVNDAP 996
Db 62 LYYDYETGSGLDGDBDEPHIFRIFNLIDNFMGEGGNRNQNDREVLIVLLDVNDAP 121
QY 997 M 997
Db 122 L 122

RESULT 11

US-08-326-117B-16
Sequence 16 Application US/08326117B
Patent No. 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-326-117B-16

Query Match 4.2%; Score 376; DB 1; Length 120;
Best Local Similarity 65.6%; Pred. No. 1.3e-24;
Matches 80; Conservative 12; Mismatches 28; Indels 2; Gaps 2;

QY 1116 VFPPQDSVIRLSRRATEGGLATPANEFLPIYATDEDDGLHAGSVTFHVGNEAVOYF 1175
Db 1 VFPPNDAVIRLARERAVINGLATVNGEFLERISATDPDGLHAGVTFQY-GDESSQRYF 59
QY 1176 DITEVGAGENGSQLILNLPFEOIRORPITRATDGGTEPGPLMTDVTFSVVFPTGGDP 1235
Db 60 QVVD-NNGENIGSLRLQAVPEELRFRITRATDGGTDPGLSTDMTFRVVFPVTOGEP 118
QY 1236 VF 1237
Db 119 RF 120

RESULT 12
US-08-982-129-16

Sequence 16 Application US/08982129
Patent No. 6007981
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,129
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/326,117
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-982-129-16

Query Match 4.2%; Score 376; DB 3; Length 120;
Best Local Similarity 65.6%; Pred. No. 1.3e-24;
Matches 80; Conservative 12; Mismatches 28; Indels 2; Gaps 2;

QY 1116 VFPPQDSVIRLSRRATEGGLATPANEFLPIYATDEDDGLHAGSVTFHVGNEAVOYF 1175
Db 1 VFPPNDAVIRLARERAVINGLATVNGEFLERISATDPDGLHAGVTFQY-GDESSQRYF 59
QY 1176 DITEVGAGENGSQLILNLPFEOIRORPITRATDGGTEPGPLMTDVTFSVVFPTGGDP 1235
Db 60 QVVD-NNGENIGSLRLQAVPEELRFRITRATDGGTDPGLSTDMTFRVVFPVTOGEP 118
QY 1236 VF 1237
Db 119 RF 120

RESULT 13
US-08-326-117B-8
Sequence 8 Application US/08326117B
Patent No. 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER

Query Match 4.0%; Score 352; DB 1; Length 124;
 Best Local Similarity 56.2%; Pred. No. 1.6e-22;
 Matches 68; Conservative 17; Mismatches 36; Indels 0; Gaps 0;
 QY 755 GTLTQTMRYRELASGTLIGSVLATDIDGPLYNQVRYTIQPRNNPEGLVKIDFTTGQIE 814
 ||| | |||:| | :||| | ||| ||| ||| | : | : | : |
 Db 3 GTLEQNFRVREMSAGGLVGSVRADDIDGPLYNQVRYTIFPREDDTKDLIMIELPHGSNF 62
 QY 815 VDANEALIDDEPWRFYLYTYVIADECSELENRTCEPPDSNYFEVPGDIELEITDTNNKVP 874
 ||| | |||:| | :||| | ||| ||| ||| | : | : | : |
 Db 63 REHKRIDANTPPRHLYTYVVASDRGSTEDPADCPPDPYMETEGNTITHITDINNKP 122
 QY 875 E 875
 Db 123 Q 123

Search completed: January 15, 2002, 17:20:45
 Job Name: 4803 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:12:57 ; Search time 108.48 Seconds
(without alignments)
2315.170 Million cell updates/sec

Title: US-09-715-909-2

Perfect score: 8909
Sequence: 1 MGVERFFPAVLVSLASAL.....THGNNGFKSPYLPQHPK 1717

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

hed: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:REMBL_17.*
2: SP:archaea.*
3: SP:bacteria.*
4: SP:fungi.*
5: SP:human.*
6: SP:invertebrate.*
7: SP:mammal.*
8: SP:mhc.*
9: SP:organelle.*
10: SP:plant.*
11: SP:rodent.*
12: SP:virus.*
13: SP:vertebrate.*
14: SP:unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5607.5	62.9	1715	09XY09	09XY09 bombyx mori
2	5597.5	62.8	1716	09NDN5	09ndn5 bombyx mori
3	5595.5	62.8	1716	09NDN4	09ndn4 bombyx mori
4	5594.5	62.8	1716	09NDN3	09ndn3 bombyx mori
5	5429.5	55.8	1717	09GPJ9	09gpj9 manduca sex
6	4631	52.0	1728	09P137	09p137 manduca sex
7	1039.5	11.7	2005	09VEH5	09veh5 drosophila
8	818.5	9.2	3322	099NH1	099nh1 mus musculu
9	818.5	9.2	3354	099PE4	099pe4 mus musculu
10	758.5	8.5	2552	09H4K9	09h4k9 homo sapien
11	758.5	8.5	3354	09H2S1	09h2s1 homo sapien
12	622	7.0	1374	09VGS5	09vgs5 drosophila
13	616.5	6.9	3503	024292	024292 drosophila
14	595	6.7	1820	09VVG0	09vvg0 drosophila
15	593	6.7	4587	09QXA3	09qxa3 mus musculu
16	585	6.6	4349	09VY08	09vy08 homo sapien
17	579.5	6.5	4643	09VW71	09vw71 drosophila
18	574	6.4	4589	09WU10	09wu10 rattus norv
19	560	6.3	4351	088277	088277 rattus norv

20	530	5.9	3579	5	09V5N8	09v5n8 drosophila
21	519	5.8	2923	4	09HCU4	09hcu4 homo sapien
22	503	5.6	2920	11	09R0M0	09r0m0 mus musculu
23	498.5	5.6	2809	5	061230	061230 lytechinus
24	494.5	5.6	3312	4	09NY07	09ny07 homo sapien
25	492	5.5	1311	5	09TYW6	09tyw6 caenorhabd
26	484.5	5.4	1439	5	09XWM6	09xwm6 caenorhabd
27	483.5	5.4	3313	11	088278	088278 rattus norv
28	452	5.1	3034	11	035161	035161 mus musculu
29	448.5	5.0	1943	11	099PJ1	099pj1 mus musculu
30	444.5	5.0	3014	4	09NY06	09ny06 homo sapien
31	442.5	5.0	1310	4	09BYE9	09bye9 homo sapien
32	442	5.0	1181	4	09Y526	09y526 homo sapien
33	434.5	4.9	4307	5	019319	019319 caenorhabd
34	405	4.5	2163	5	001912	001912 caenorhabd
35	403	4.5	2610	5	019482	019482 caenorhabd
36	374.5	4.2	3337	5	09TWY4	09twy4 caenorhabd
37	369.5	4.1	1965	2	P73089	P73089 synecocyst
38	366.5	4.1	2204	5	09VEU1	09veu1 drosophila
39	353	4.0	2240	5	09GR86	09gr86 drosophila
40	352.5	4.0	1340	4	09B2A8	09b2a8 homo sapien
41	352	4.0	1563	5	09VAF5	09vaf5 drosophila
42	344.5	3.9	829	4	09Y5D4	09y5d4 homo sapien
43	344.5	3.9	932	4	09Y5H0	09y5h0 homo sapien
44	344	3.9	1203	4	09HC56	09hc56 homo sapien
45	341.5	3.8	1337	4	09B2A6	09b2a6 homo sapien

ALIGNMENTS

RESULT 1
ID 09XY09 PRELIMINARY: PRT: 1715 AA.
AC 09XY09;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BTR175 PRECURSOR.
GN BTR175.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylsia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MIDGUT;
RA Nagamatsu Y.;
RT "Cloning, Sequencing, and Expression of the Bombyx mori Receptor for
RT Bacillus thuringiensis Insecticidal CryII(a) Toxin.";
RT Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC TISSUE=MIDGUT;
RA Nagamatsu Y.;
RT "Cloning, Sequencing, and Expression of the Bombyx mori Receptor for
RT Bacillus thuringiensis Insecticidal CryII(a) Toxin.";
RT Biocli. Biotechnol. Biochem. 62:727-734(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO THE CADHERIN FAMILY.
DR EMBL, AB026260; BAA77212.1; .
DR HSP, P15116; INCI.
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 6.
DR SMART: SM00112; CA; 10.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS50268; CADHERIN_2; 10.
KW Calcium-binding; Cell adhesion; Glycoprotein; Repeat; Signal;
KW Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 1715 AA; 19335 MW; 2EC68401FB054C0F CRC64;

Query Match 62.9% Score 5607.5; DB 5; Length 1715;
 Best Local Similarity 63.5%; Pred. No. 0;
 Matches 1095; Conservative 224; Mismatches 377; Indels 29; Gaps 12;

1 MGVEFPFVALLVSLASLANQRCSYIIAIPRPETPELPIDYEGKSNSEOLIPGPT 60
 1 MGVDRIILATLLIYAETVLAQERGFVWALPRPRPDLPELDFSGQWMSQRLIPADR 60
 61 EYVCMENF---LPDMQIYIWEEBIEGDYIIAKLNYGSGNTPLV--SLMSGOPRAQLCP 115
 61 EDVCMGDGHAMPPTGTGTQIYWEEBIEGEVPLAKLNYRPNVYIEPVLSSGFNLVAV 120
 116 EPRQWAGQMSLVYTORQDYETATMOSVFSIOEGESQAVLVALEYINIDNPILQV 175
 121 IRRIPDSNGEMHLITQKODYETPGMOQYVFNIRIDGETLVAGVSLIYVNDNAPIIDA 180
 176 VSACVYPEHGAEARLDYQVYSDRDEGEISTREPMTERFVDSRAADESIFVMEYDPSDMF 235
 181 LEPQGVDELGEARLFEQYVYVTDADGRISTQFMQFRIDSR-GDCKIFYIGCANIPGEMI 239
 236 NKKMTVGINSPINETTQHLIFSVTASDSLPHNNHTVMMVQYVENESRPKRVVEIFSVQ 295
 240 RMTMTVGINELPNETNPLHIFSVALDSLPTHTVTLLVQYVENEHRRPVRWEIIFAVQ 299
 296 PEKTNOSPLRAIDGDIINRAINVTLLIRDADDFSELEVE--DGAHLVTEIDRX 352
 300 PEKTAQSPVPAIDGDIINRAINVTLLIRDADDFSELEVE--DGAHLVTEIDRX 359
 353 LERELFNLTIVAKSTDAFSAFETAHFIIVNDVNDQRPRLKHEYSIDIMEETPLNFN 412
 360 LQREVFOGLIAYKXKDNSSATANVLIIVNDINDQRPRLKHEYSIDIMEETPLNFN 419
 413 EBFQGHDRDLGNAQYVELEDFEPPGASAFYIAPGSGYQOTFMGINTMTLDYEV 472
 420 QFEGHHRDLGNAQYVELEDFEPPGASAFYIAPGSGYQOTFMGINTMTLDYEV 479
 473 IFQNTIIVKAVDMNNAHSVGEALVYVNLINNDDELPIEESYSASFKEYGAGPVPAT 532
 480 EFQRIIRKRIATDMNEEVAAYVYIINLINNDDEPIEHSVQVNSFETEGKGFVAN 539
 533 VLALDRIDIDVYVSLMGNAVDFLEFIDESTGEIVYSMDADFYHRONTLFEVORADDTLG 592
 540 VYAHNRDIDDRVEHETLMGNAANYLSIDKDTGDIHVTDFFYHROSLEFLVQVRADDTLG 599
 593 DQPHNTVTLQVLEIEDVNNPTPLRLPRSTPSVEENVEGEVISELITATDPDTSAYLM 652
 600 E-PEHPATSQLIHLIEDVNNPTPLRLPRGSPVNEENVEGEVITITSEIRATDPDTSAYLM 658
 653 FEIDMSFWATKQGRETNPTEYVGCIVETIYPTTEGNRGSAGRLVAQVEIRDNVTIDFEE 712
 659 FEIDMTTSVATQGRANPIEFHNCVEIETIYPALNNRGSAGRLVVKIRREVITIDFEE 718
 713 FEMLYLTKVWRLDNLTVIGDYDEAFETIITIDMDNAPIFANGTILQTMKRVRELAASGLT 772
 719 FEMLYLTKVWRLDNLTVIGDYDEAFETIITIDMDNAPIFANGTILQTMKRVRELAASGLT 778
 773 TGSVATADIDGLYQVYKTIQPRNNTPEGLVKIDFTTQOIEVDANEAIDADEPFAFYLY 832
 779 IGTLTATDIDGLYQVYKTIQPRNNTPEGLVKIDFTTQOIEVDANEAIDADEPFAFYLY 838
 833 YVIVASDECSLNRRECPDPSNVFEVPGDIEIETIDTNKVPPEPLTEKNTFYVYVENAT 892
 839 YVIVATDRCTAEDPDDCPDPTWETPGOVVIOIITDNKIKQPEIDQKAVVYITIEDAV 898
 893 SGDEVVQVLSHRDRDELHYTVRYTMFAVNPRLDFEVDLDTGRLLEVYPGDEKLDR 952
 899 SGDEVVQVLSHRDRDELHYTVRYTMFAVNPRLDFEVDLDTGRLLEVYPGDEKLDR 958
 953 GDEPHTITIVNFDNFSQDGRANDVEITVLLDVNDNAPEMALPDELPRDVSSEGA 1012
 959 GDEPHTITIVNFDNFSQDGRANDVEITVLLDVNDNAPEMALPDELPRDVSSEGA 1018

1013 AGRVLPETIYAPDRDEPPTDNRVYGIILDTITDRDIEVPLFTMISLENKTEGETAM 1072
 1019 KTRIQPHIYAPDRDEPPTDNRVYGIILDTITDRDIEVPLFTMISLENKTEGETAM 1078
 1073 DIRGYWYEIETEARPDHGYPOORSNEYTYLVIRPNHHVYVFPQDPSVRLSREAR 1132
 1079 DIRGYWYEIETEARPDHGYPOORSNEYTYLVIRPNHHVYVFPQDPSVRLSREAR 1138
 1133 EGVGLATANAEFLPELVYATDEGLHAGSVTFPHVQCNNEAVOYEDTTEVAGGNSGOLIR 1192
 1139 VNGLATYVDEGLHAGSVTFPHVQCNNEAVOYEDTTEVAGGNSGOLIR 1196
 1193 QLEPEIQRFITIRATIGTEPGLMTDVTFPSVYVFPQDPSVSENAATVAFEGEGC 1252
 1197 QLEPEIQRFITIRATIGTEPGLMTDVTFPSVYVFPQDPSVSENAATVAFEGEGC 1256
 1253 LRESELPQAEIDLKHLCEDCODIYFRFDGNNEGFLVLDQSSNVISLAOLDEEVATS 1312
 1257 LRESELPQAEIDLKHLCEDCODIYFRFDGNNEGFLVLDQSSNVISLAOLDEEVATS 1316
 1313 YTLHTAASNSPATGIPLOTSTLVTVNVRANPRPIEBODLYTAGISPLSISGRELTV 1372
 1317 YTLHTAASNSPATGIPLOTSTLVTVNVRANPRPIEBODLYTAGISPLSISGRELTV 1374
 1373 RASHTEDDTITTYIDRASQMLDSLEAVDSAFALHATGVLSLNMQPTASMHGFEFVY 1432
 1375 QATHSENAPIITITDMSVMTVPTLASVRETAFLINPHGVTLINQPTASMHGFEFVY 1434
 1433 IATDTRASADTRAVKYLISQNRVTFIFDNOLETVEONRNFIAATFSGFMNTCIDOV 1492
 1435 VATDPAGYSDRANVXYLSTSTRNVEFLVNTLEQVEQNTDEIAQTFSGFMNTCIDOV 1494
 1493 VPESSSGVAODTEVRHAFIRDNVYVQAEVAVRSQTVLRTQLMLSTNSLVLOL 1552
 1495 VPESSSGVAODTEVRHAFIRDNVYVQAEVAVRSQTVLRTQLMLSTNSLVLOL 1554
 1553 VTGDTPTLGEESQIYVLAALSAVGLFCLVLLALEPCRTALNRQALSMRTKYSV 1612
 1555 VTGDTPTLGEESQIYVLAALSAVGLFCLVLLALEPCRTALNRQALSMRTKYSV 1613
 1613 DGLNRAGL-AGCTKHAHVEGSPKMKNEAIRAPDPAIDASGDSPLGIEIMPQORDY 1671
 1614 DGLNRAGL-AGCTKHAHVEGSPKMKNEAIRAPDPAIDASGDSPLGIEIMPQORDY 1673
 1672 FPPGDTSSSGIYVLMHGEATDNKP---VTHGNGFGKSTPLYPO 1713
 1674 FPPGDTSSSGIYVLMHGEATDNKP---VTHGNGFGKSTPLYPO 1708

RESULT 2
 09AND5 PRELIMINARY; PRT: 1716 AA.
 AC 09AND5;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE CADHERIN-LIKE MEMBRANE PROTEIN (FRAGMENT).
 GN BTR175A.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MIDGUT;
 RA Ikawa S., Tsuda Y., Fukada T., Sugimoto K., Himeeno M.,
 RT "cDNA cloning of the CryIa receptor variants from Bombyx mori and
 RL their expression in mammalian cells.";
 CC Submitted (Apr-2000) to the EMBL/Genbank/DBD databases.
 DR EMBL; AB041508; BAA99404.1; -.

RP SEQUENCE FROM N.A.
 RC TISSUE=MIDGUT.
 RA Ikawa S., Tsuda Y., Fukada T., Sugimoto K., Himeno M.:
 RT "cDNA cloning of the Cytlaa receptor variants from Bombyx mori and
 RL their expression in mammalian cells.";
 CC Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR -1- SIMILARITY: TO THE CADHERIN FAMILY.
 DR EMBL: AB041509; BAA99405.1; -
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin; 6.
 DR SMART: SM00112; CA; 10.
 DR PROSITE: PS00232; CADHERIN_1; 1.
 DR PROSITE: PS00268; CADHERIN_2; 10.
 DR Calcium-binding; Cell adhesion; Glycoprotein.
 KW NON_TER 1716 1716
 SO SEQUENCE 1716 AA; 193323 MW; 0C2D5C3571097BC8 CRC64;

Query Match 62.8%; Score 5595.5; DB 5; Length 1716;
 Best Local Similarity 63.4%; Pred. No. 0;
 Matches 1093; Conservative 224; Mismatches 379; Indels 29; Gaps 12;

1 MGVEFFPFAVLVLSAALANORCSYIAIPREPTPELPIDYEGKMSQGLIPGPT 60
 1 MGVDRIIATLLIYAEVTLAQRGCFMVAIPRPRPDLPEIDFEGQWTWSORPLIPADR 60
 61 EGVCEMF---LPDQIYIYMEEEIEGDIYIAKLYOGSNTPLV--SMSGOPRAQLCP 115
 61 EDVCGDYGHAMTPYGTGIIYMEEEIEGDIYIAKLYOGSNTPLV--SMSGOPRAQLCP 120
 116 EFRONEADQMSLVITORYDETATMOSYVFSIQVSGESQAVLALEIVINDDNPILQV 175
 121 IRRIPDSNGEHLITITQODYETPMQOYVFNIRIDGETLVAGVSLIVINDDNPITQA 180
 176 VSACVIRPHEGARLDYVYOSDRGELSTRPMTFRVDSRRADISYVMEYSPDMF 235
 181 LEPCCVDLGEARLRECYVVTADGRISTQPMQFRIDSDR-GDCKIFYIQGANITGEVI 239
 236 NMKATVGINSPINFEETTLHIFSVTASDLSLNNHTVTMMVQVENESEPPRAVEIFSVQ 295
 240 RMTVMVGINPELFTNPLHIFSVTALSDLSLNNHTVTMMVQVENESEPPRAVEIFSVQ 299
 296 FDEKTNQSFSLAIDIGDGINRAIYTLIRDDADPFSLVYE--DGAIIYVEIDRD 352
 300 FDEKTAQSFPRALIDGDGINKPIHYRLETAEDPFHRIITEGGRSAILVDPIDRPT 359
 353 LERELFNTIYAYKSTASPAETEAHFIIVNDVNDORPEPLHKEYSIDIMEETPTMUNP 412
 360 LQREVFSLIATKIDNESSATIAANVAVIIVNDIDOREPEPLKEYRLINIMESTALTLPND 419
 413 EEEGFHDDLGENAQYVTELEDVPPPGASAFYIAPGSGYORQFTINGTINTMLDYEDV 472
 420 QEEGFHDDLGENAQYVTELEDVPPPGASAFYIAPGSGYORQFTINGTINTMLDYEDV 479
 473 IFQNTIYKAVADNMNASHVEGALYVNLINMNDLPTFEESYSASKEVYAGAPPAVAT 532
 480 EFORIRLEIVITIDMNEHGVAYAYVYINLINMNDDEPIFEHSVQVNSFEKEGKEFVAN 539
 533 VLADRODIDVYVSHNNAVYDLFIDESTGEIYFMSMDAPYHNOHNTLFYOVARDITLG 592
 540 VRAHRODIDVYVSHNNAVYDLFIDESTGEIYFMSMDAPYHNOHNTLFYOVARDITLG 599
 593 DGPANTVTQVLELEDVNNPTPLRSPSPVSENVPEGEYREIETATPPTDSAYLM 652
 600 E-PFTATISQLLIHDDIINNPTPLRSPSPVSENVPEGEYREIETATPPTDSAYLM 658
 659 FEIDMTSYATKOGREANPIEFHNCVEIETIYPAIINNGSAIGRLVYKIRENVITIDEE 718
 713 FEMLYLTVRVARDLNTVIGDDYDEAFITTIIDMNDAPIFANGILTQYMRARELAAGTL 772
 719 FEMLYLTVRVARDLNTVIGDDYDEAFITTIIDMNDAPIFANGILTQYMRARELAAGTL 778

773 IGSVLTATIDIDGLYNQVRYTIOPRNNTPGGLVYIDFTTGQIEVDANEAIDAEPWRFY 832
 779 IGTLTATIDIDGLYNQVRYTIOPRNNTPGGLVYIDFTTGQIEVDANEAIDAEPWRFY 838
 833 YTVIASDECSLENRETCPPDSNYFEVPGDIEIETIDTNKKVPELTERKNTTVYEMANT 892
 839 YTVVATIDRSYADPDCCPDPPYMETPCGVYIQTNNKIPQEPEDOFKAVAYIYEDAV 898
 893 SGDEYVQVLSHRRDDELHYVRYTMNFAVNRLDFFFAVDPTGRVYVYITDYEDLRD 952
 899 SGDEYVQVLSHRRDDELHYVRYTMNFAVNRLDFFFAVDPTGRVYVYITDYEDLRD 958
 953 GDEPTHTYFVNEIDNFEPSDGRNRNDQEVLEFVLLDVNDNAPRLPELDEFYSEGA 1012
 959 GDEPHRIEFENLIDNFPQDQGNRNQNDQEVLEFVLLDVNDNAPRLPELDEFYSEGA 1018
 1013 AGVRLPEIYAPDRDEPDTNSRVGIGLDTITTRDIEVPRDLFTMISTENKTGLEETAM 1072
 1019 KGTRLQPHIYAPDRDEPDTNSRVGIGLDTITTRDIEVPRDLFTMISTENKTGLEETAM 1078
 1073 DLRGVWGYEIPFIEAFDGYPOQSRNETHYTYLIRPYNHHNPFVPPQDSVYRLSRERAT 1132
 1079 DLRGVWGYEIPFIEAFDGYPOQSRNETHYTYLIRPYNHHNPFVPPQDSVYRLSRERAT 1138
 1133 EGVLTATANELEPIYATDEDEGLHAGSVTFHVQNEEAQVYFDTTEVGAQENSGQLIR 1192
 1139 VNGLTATYDGEFLNRIYATDEDEGLHAGSVTFHVQNEEAQVYFDTTEVGAQENSGQLIR 1196
 1193 QLPFQIQRQRTITATDGTGTPGLMTDYTPSVYFVPLQDGPVSENAATYAFTEGEBG 1252
 1197 QLPFQIQRQRTITATDGTGTPGLMTDYTPSVYFVPLQDGPVSENAATYAFTEGEBG 1256
 1253 LRESFELPOAEDLKNHLEDDEDCODIYRFRIDGNNGELFVLDQSSNVISLAOLEBVA 1312
 1257 LRESFELPOAEDLKNHLEDDEDCODIYRFRIDGNNGELFVLDQSSNVISLAOLEBVA 1316
 1313 YTLIAAASNSPDATGIPLOTSLIVTVVYVREANPRPIEDODLYAGISTLDSIGRELLTV 1372
 1317 YTLIAAASNSPDATGIPLOTSLIVTVVYVREANPRPIEDODLYAGISTLDSIGRELLTV 1374
 1373 RASHTEDDTITITIDRASQDSSLEAVRDSAFALHATTVGLVSLMMQPRASHNGAFEDV 1432
 1375 QATSHENAPIYITIDSLSTVTPDLASVRETAFLNPHGVTLNLIQTPASHNGAFEDV 1434
 1433 IATPTASADTARVYUULISSONRYTEIFDNOLETVEQNRNFIATFSTGFMTCNIDOV 1492
 1435 VATDPAAGSDRANVYUULISSONRYTEIFDNOLETVEQNRNFIATFSTGFMTCNIDOV 1494
 1493 VPFSDSGVAODDTTEVRAHPIRDVNPVQAOEVAVRSDVTLKTIQMLSTNSVLQDL 1552
 1495 VATDPAAGSDRANVYUULISSONRYTEIFDNOLETVEQNRNFIATFSTGFMTCNIDOV 1554
 1553 VTGDPFTLGEESMQTAVALALASVLFGLCVLLALFCRRALNRQLOALSMRTKYSV 1612
 1555 FT-DTSPAPDAGSAVALALASVLFGLCVLLALFCRRALNRQLOALSMRTKYSV 1613
 1613 DSGLRNAGL-ARGTNKHAHVEGNSPMNNEAIRAPDADSDAGSDGLGIEDMPQFRDY 1671
 1614 DSGLRNAGL-ARGTNKHAHVEGNSPMNNEAIRAPDADSDAGSDGLGIEDMPQFRDY 1673
 1672 FPPGDTSSSGIIVLMGATDNKP--VTHGNNGFSTPYLPQ 1713
 1674 YFRPVD-----EFKTDKPEDYATANNNGFSTPSP 1708

RESULT 4
 Q9NDN3 PRELIMINARY; PRI: 1716 AA.
 AC Q9NDN3
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE CADHERIN-LIKE MEMBRANE PROTEIN (FRAGMENT).
 GN BRL75C.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 NCBI_TaxId=7091;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=MIDGUT;
 RA Ikawa S., Tsuda Y., Fukada T., Sugimoto K., Himeno M.,
 RT "cDNA cloning of the CryIa receptor variants from Bombyx mori and
 RL their expression in mammalian cells."
 Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE CADHERIN FAMILY.
 DR EMBL: AB041510; BAA99406.1; -;
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin; 6.
 SMART: SM00112; CA; 10.
 PROSITE: PS00232; CADHERIN_1; 1.
 DR PROSITE: PS50268; CADHERIN_2; 10.
 KM Calcium-binding, cell adhesion; Glycoprotein.
 FT NON-TER 1716 1716
 SQ SEQUENCE 1716 AA; 193396 MW; 7A8151227C15C43A CRC64;

Query Match 62.8%; Score 5594.5; DB 5; Length 1716;
 Best Local Similarity 63.4%; Pred. No. 0;
 Matches 1093; Conservative 224; Mismatches 379; Indels 29; Gaps 12;

QY 1 MGEREFPALVILSLASAAIANORCSYIIAPREPPELPIDIECKSMSEOPPLIGPTR 60
 DB 1 MGVADVILATLILIVAEYLAQRCGFMAIPRPRLDELDELDEGCTMSQRLLIAPADR 60
 QY 61 EYECMENF---LPDQMIQVITMEETIEGDIYAKLVQSSNPVL--SIWGGPRAALGP 115
 DB 61 EDVCMQGYNHMPPTGYTQIIVMEETIEGEVPIAKLVNRCGNVPIEPAFLSGFNLLVPI 120
 QY 116 ERFQNRADGQMSLVITORDYETATMOSYVFSIQVGEQANVLALEYINIDNPILQY 175
 DB 121 IRIIPDSNGMLITLITQORDYETPGMOQYFNIRIGETILVAGVSLIYINIDNAPILQ 180
 QY 176 VSACVIREGEALRLDCYQVSDRDEISTREMTFRVDSRADESIFVMGVEYDSDMF 235
 DB 181 LERQYDELGEALTECYVYVTDADGRISTQFWQFRIDSDR-GDKIIFITQAGNIGEMT 239
 QY 236 NKMATVGINSPLEFETTOIHISVYASDSLPHNHVYTMVQVENVESRPPRWEIFSVQO 295
 DB 240 RMTMTGIMEPLNFETNPLHISVYALDSLPNHTVTLMVQVENVESRPPRWEIFSVQO 299
 QY 296 PDEKTNQSLRADGDDGQINRAINTLIRADADDFEFLSEVIE--DGAIIHYTEIDRDK 352
 DB 300 FDEKTAQSPVRAIDGDTGINKIRHLETAEDYFHLRTLEGSGALILVDPIDRDT 359
 QY 353 LERLENLIVAYKSTDAFATEAHIFIIYVNDQDRPEPLHKEYSIDIMEETPMLTLEN 412
 DB 360 LQREYVQLSIAYKYVNDSSATAANVITVINDQDRPEPLKEYVNLNIMEETALFLND 419
 QY 413 EERGFHRLDQLEAQTVELEDFPFGAASAFYIAGSSYQOROTFMGTINTMLDYEDV 472
 DB 420 QEGFHDRLDQGAQVYVLESYDPRADAKAFAVEYQOROTFMGTINTMLDYEDV 479
 QY 473 IFQNTIIRKAVDMNNAHSVGEALYVNLINNNDELPIFEESYSASEKETGAGAPVAT 532
 DB 480 EFQRIKLRIATDMDEEYVGAIVYINLINNDEEPIEHSQVNSFEKETGEGKFEVAN 539
 QY 533 VIALRDIDVYVHSLMGNAVYDLFIDESTGELFVSMDDAFDYHRONTLFYQVRADTLG 592
 DB 540 VRAHRDIDDRVEHITLMGNANNVYLSIDKOTGDIHYQDDFPYHRSSELYQVRADDLIG 599
 QY 593 DGRPHNVITQVILEDVVNTPTLRLPRSTPSVEENVBEGEISEIETATDPRTSAYLM 652
 DB 600 E--PFIHTATSQLIHLDDINNTPTLRLPRGSPNVEENVBEGYIITSEIRATDPDTAELR 658

QY 653 FEIDMDSTWATKQGRENPTEYGCIVETIYPTEGNRGSAIGRLVQOIRDNVTIDFEE 712
 DB 659 FEIDMTTSYATKQGRANPLEFNANCVIEITTYAIRNNRGSALIRLVYKATRENVITDYE 718
 QY 713 FEMLVLTVRVRLNTVIGDDYDEATFTITIDNDNAPIFPANGTLQTMKRELAASGTL 772
 DB 719 FEMLVLTVRVRLNTVIGDDYDEATFTITIDNDNAPIFPANGTLQTMKRELAASGTL 778
 QY 773 IGSVLAFTDIDGPELVNOVRTIOPRNNTPREGVLKIDTGTQIEVDANEAIDADPRFYLX 832
 DB 779 IGLTFTDIDGPELVNOVRTIOPRNNTPREGVLKIDTGTQIEVDANEAIDADPRFYLX 838
 QY 833 YTVIADDESLERNKCPDPSNFVEVPGDEIEIIDTNKVPPELPEKFTYVYENAT 892
 DB 839 YTVIADDESLERNKCPDPSNFVEVPGDEIEIIDTNKVPPELPEKFTYVYENAT 898
 QY 893 SGDEVVQLYSHDRDELHYTVRYTNNFAVNPRLRDFEVDLDTGRLEHYHYPGDEKLRD 952
 DB 899 SGDEVVQLYSHDRDELHYTVRYTNNFAVNPRLRDFEVDLDTGRLEHYHYPGDEKLRD 958
 QY 953 GDEPTFTIYVNFIDNFFSDGGRNODVEIEVYLLDVNDNAPEMLPDELRFVSEGAH 1012
 DB 959 GDEPTFTIYVNFIDNFFSDGGRNODVEIEVYLLDVNDNAPEMLPDELRFVSEGAH 1018
 QY 1013 AGYRVLPEIYAPDRDPRDIDNSRVYGIILDTITDIDIEVDPDFTMISTENKTELETAM 1072
 DB 1019 KGRILDPHIYAPDRDPRDIDNSRVYGIILDTITDIDIEVDPDFTMISTENKTELETAM 1078
 QY 1073 DLRGVGTGTEIFEAEDHGYPOORSNETYLVIRPNFHPVFPQPSVIRLSERAT 1132
 DB 1079 DLRGVGTGTEIFEAEDHGYPOORSNETYLVIRPNFHPVFPQPSVIRLSERAT 1138
 QY 1133 EGGVLAFAANLEPIYATDEDEGLHAGSVTFHVQNEAAYQVFDITEVAGENSGLILR 1192
 DB 1139 VNGILATVGEFENIRIYATDEDEGLHAGSVTFHVQNEAAYQVFDITEVAGENSGLILR 1196
 QY 1193 QLEPQIRQFRTIRATDGTGTEPMTDVTFSVYVPPGQDVPFSENAATYAFEGEG 1252
 DB 1197 QLEPQIRQFRTIRATDGTGTEPMTDVTFSVYVPPGQDVPFSENAATYAFEGEG 1256
 QY 1253 LRESFELPOAEDLKNHLCEDDCDIYREFIDGNEGLFVLDOSSNYISIAOELREVATS 1312
 DB 1257 LRESFELPOAEDLKNHLCEDDCDIYREFIDGNEGLFVLDOSSNYISIAOELREVATS 1316
 QY 1313 YTLHTAASNPDAATGIPLOTSLVTVVNVREANRPRIEODLTAGTISTLDSIGRELLTV 1372
 DB 1317 YTLHTAASNPDAATGIPLOTSLVTVVNVREANRPRIEODLTAGTISTLDSIGRELLTV 1374
 QY 1373 RASHTEDDTITTYIDRASQDLSLEAVRDSAPALHATGVLISLNNQPAASHMGFEFV 1432
 DB 1375 RASHTEDDTITTYIDRASQDLSLEAVRDSAPALHATGVLISLNNQPAASHMGFEFV 1434
 QY 1433 IATDPAASIDTARVKKVYLISQNRVYFIDNQLETEYQNRNFIAFTSTGFNMTCNDIV 1492
 DB 1435 IATDPAASIDTARVKKVYLISQNRVYFIDNQLETEYQNRNFIAFTSTGFNMTCNDIV 1494
 QY 1493 VAPDSAGYDDBRANVKKYLISQNRVYFIDNQLETEYQNRNFIAFTSTGFNMTCNDIV 1552
 DB 1495 VAPDSAGYDDBRANVKKYLISQNRVYFIDNQLETEYQNRNFIAFTSTGFNMTCNDIV 1554
 QY 1553 VAPDSAGYDDBRANVKKYLISQNRVYFIDNQLETEYQNRNFIAFTSTGFNMTCNDIV 1612
 DB 1555 VAPDSAGYDDBRANVKKYLISQNRVYFIDNQLETEYQNRNFIAFTSTGFNMTCNDIV 1613
 QY 1613 DSGLNKAGL-APGTNKHAVEGSPNMNNEAIRAPDEDAISDAGSDSLGIEDMPQRFDY 1671
 DB 1614 DSGLNKAGL-APGTNKHAVEGSPNMNNEAIRAPDEDAISDAGSDSLGIEDMPQRFDY 1673
 QY 1672 FPGGDTSSSGVILMGEATDNKP---YTHGNNGEFTPLPQ 1713
 DB 1674 FPGGDTSSSGVILMGEATDNKP---YTHGNNGEFTPLPQ 1718

RESULT 5
ID 096P39 PRELIMINARY: PRT: 1717 AA.
AC 096P39;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CADHERIN-RELATED PROTEIN RECEPTOR BR-R1.
OC Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_Taxid=7130:
RN [1]
RP SEQUENCE FROM N.A.
RA Dorsch J.A., Maaty W.S.A., Griko N.B., Candas M., Bulla L.A., Jr.;
RT "A Cadherin-related Protein Receptor, BR-R1, in the Midgut Epithelium
of Manduca sexta Mediates Toxicity for Bacillus thuringiensis CryIA
Toxins."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: TO THE CADHERIN FAMILY.
EMBL, AF319973, AAC37912.1;
InterPro: IPR001525; C3_DNA_meth.
DR InterPro: IPR00126; Cadherin.
DR Pfam: PF00028; cadherin; 6.
DR SMART: SM00112; CA; 10.
DR PROSITE: PS00095; C5_MWASE_2; UNKNOWN_1.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS00268; CADHERIN_2; 10.
KW Calcium-binding; Cell adhesion; Glycoprotein; Receptor.
SQ SEQUENCE 1717 AA: 192305 MW: FEC4A48B098B17E CRC64;

Query Match 59.88: Score 5329.5; DB 5; Length 1117
Best Local Similarity 60.28: Pred. No. 0;
Matches 1049; Conservative 249; Mismatches 396; Indels 33; Gaps 16;

QY 1 MGVEFFPAVLIVSLASALANORCSYIIAIPR-PEPELPIDYEGKSMSEPLIPGPT 59
DB 1 MAVDVRIAPFLVFLAPLAVOERCGYMTAIPRLPRPNLPVLANEGOTWSGRPLPAPE 60
QY 60 REEVCMEFNE---LPDQMLQVYMEEEIEGDVIAKLNTYQGSNTPLSLI--MSGPRAOLG 114
DB 61 RDLCLMDAYHVTANLGTQVYIMDEIDEITAILNTNGPSTPIELPELSGSYNL-LM 119
QY 115 PEFQNEADGQMSLVITRODIETATMOSYVPSIOVEGSOAVLVALEIVNIDNPILQ 174
DB 120 PVIRFVD-NGEMHLITIRORHYELPGMOQYMFNVVDSQSLVAGVSLAIVNIDNAPILQ 178
QY 175 VVASCVIPEHGRARLTDCVVOYSDRDGEISTRFMFVRVDSRAADESIFVWVGEXDPSDW 234
DB 179 NFEPCRVLELGPGLTECTTQVSDADGRISTEFMTFRIDSVR-GDEETFYIERTINPNOW 237
QY 235 FNMKATVGINSLPNETQTLHFTVYASDSLPHNHTVTMNOVENVESRPWEVIFSVQ 294
DB 238 MMLNNTIGVNTSLNVTSPHIFSVTALDSLPHHTVTMMQVAVNSRPPWLEIFAVQ 297
QY 295 QPEDTNSQSFIRALDGTGINKRAINTYLRDADDFSLSEVI---EDGAILHVEIDRD 351
DB 298 QFEESYQNFYVIRALDGTETINMPINRYLRINEEDTFFSIEALPGKSGAVFLVSPIDB 357
QY 352 KLERLFLNLTIAVYKSTASPFATFAHIFIVNDVNDORPEPLHKYSIDIMEETPMTLNF 411
DB 358 TLQREVFLITVAYKYDEFASTSTNVITVITDINDQRPPEPHKEVRLAIMEETPLTLNF 417
QY 412 NEEGFHRRDLGEGNAQVYELDEVPFPGASAFAFYIAPSGVQROTFTINGTINHTMLDYED 471
DB 418 DKEFEFHDKDQGNQAYIVRLSVDPGAAFAFYIAPEGYQROFTINGTINHTMLDYED 477
QY 472 VIFQNIITIKVAVDMNASHGEGALVYVNLINWDELPIFEESSTYSASFETVAGGFVYA 531
DB 478 PEFOSITIRVATDNNDRHGVAVLHIDLINWDEOPIFEHAQVOTVFDEDEGEFFVA 537

QY 532 TVLADRLDDVVVHSLGNNAVDFLEIDSTGEJFVSMDDAFDYHRQNTLFVORADDTL 591
DB 538 KAVANDRIGVVEHTLLGNVNFLLTKLIGDINVASANDSNVRESELFVQVATDTL 597
QY 592 GDGPHNVTTQVLEIEDVNNTPPTLRPRSTPSVEENVPGCEISRRITATDPTSLV 651
DB 598 GE-PHRTSOLVIRLANDINNPTLRPRSGPQVEENVPDHVTQELRATDPTTADL 656
QY 652 WFEIDMSTMTKQGRFNPTVEYVCCIVETIETPEGRGSAIGLVQOETRDNTIPE 711
DB 657 REEIMWDSFPAKQGRQANPDEFRCVETETETPEINNGLAGVAVARELNHNTIYE 716
QY 712 EEFMLYTVRVDLTNVTGDDYDEATFTITIDMDNAPIFANGATLTQMRVRELAAGT 771
DB 717 EFEVSLTVRVRADLTNYGDDYDESMITITIDMDNAPVWVEGTLQGNFRKESAGGL 776
QY 772 LIGSVLATIDGPLYNOYRYTIQPPNNTPPEGVLKIDFTTGGILEVDANEAIDADEWRYL 831
DB 777 VVGSVRADDIDGPLYNOYRYTIFPREDDKDLIMIDFLTGQISVNTSADIDATPPRHL 836
QY 832 YTVTASDECSLENTBOPDSNFEVPGDIEIETIIDNNKVPEDTEKFNFTYVWENA 891
DB 837 YTVVADRCSTEDPADCPDPTTETEGNITITIDNNKVPQAEITKFDIVYIYENA 896
QY 892 TSGDEVOLYSHDRDDELVTYRTMFAVNPRLDEFEVDLTGRLEVHP---GDEK 948
DB 897 TRLDEVVTLIASDLDRLDEIHTVSIVYINAVNPRLNFEFSVRETGLVYVDYETQSGSEV 956
QY 949 LDRDDEPHHTITFVNFIDNFSGGRNODVELEFVLLVDVNDAPPEPLDELRFVYS 1008
DB 957 LDRDDEPHHTITFVNFIDNFSGGRNODVELEFVLLVDVNDAPPEPLDELRFVYS 1016
QY 1009 BGAVAGVRLPEIYAPDRDEPTDNSRVGYGLDITLIDRDLEVPDLFTMISENKTGEL 1068
DB 1017 ENLKQGVRLPEIYAPDRDEPTDNSRVGYGLDITLIDRDLEVPDLFTMISENKTGEL 1075
QY 1069 ETAMDLRGVYTEIFTEAFDGYQOORSNETYTLVIRPNHAFVFPPOPDSVIRLSR 1128
DB 1076 ETAMDLRGVYTEIFTEAFDGYQOORSNETYTLVIRPNHAFVFPPOPDSVIRLSR 1135
QY 1129 ERATGCVLATPANEFLPIYATDEDCGLHAGSVTFHVQNEAAYQFDITTEVAGAEANSQ 1188
DB 1136 ERAVINGVATATNGEFLERISATIDPDGLHAGVTFVQVDEDSQRFQV--VNDENLGS 1193
QY 1189 LILRLQFROIROPFLTRATDGTGTEPGPLMTDVFESVYFVTOGDPVSENAATVAFPE 1248
DB 1194 LRLDAVPEIREFETTRATDGTGTEPGPLMTDVFESVYFVTOGDPVSENAATVAFPE 1253
QY 1249 GEEGRRESFELPQAEELKNNLCEDCODIYRFIDGNNGELFVLDOSSNVISLAELDRE 1308
DB 1254 KSGAMEESHQPLADKINHLCEDCODHSIYRIIDGNSGCHGLDPVNRRLKLKELLE 1313
QY 1309 VATSTLHTAASNSPDANGIPLOTSLVTVVAVREANPPIEODLYTAGISTDSIGRE 1368
DB 1314 OSASHTLQVAAANSPPD-GIPLPASLITVTVVREADPPVRELYTAGISTDSIGRE 1372
QY 1369 LITVRASTHEDDTIYTTIDRASMOELSLVAVRDSAPALHATTVGLSLNOMOTASMHGF 1428
DB 1373 LIRLHATOSGSAITVYADYDMVDPVSLVAVRDSAPALHATTVGLSLNOMOTASMHGF 1432
QY 1429 EFDVATIDTASADITARVAVYLISSONRYTEFIDNOLEVEONRNFIAFSTGFMNMCN 1488
DB 1433 KFEVATIDTASADITARVAVYLISSONRYTEFIDNOLEVEONRNFIAFSTGFMNMCN 1492
QY 1489 IDQVVPFSDS-SGVAODTTEVRAHPIRDNVVOQAEVAVSDVYLRTIOLMSTNSL 1547
DB 1493 IDQVVPANPVTGVALHESHQKRGHPIRDNVVYLADEIOISDVLVLSIOTTLAARSL 1552
QY 1548 VLODVLGTPTPLGSESNQIAYVALAASVGLFIVLILLFQRTLRNLQOLALSMF 1607
DB 1553 VLODVLGTPTPLGSESNQIAYVALAASVGLFIVLILLFQRTLRNLQOLALSMF 1611

QY 1608 KYGSDVSGINRAGL-APGTNKHAEVGSNPMNMNEAIRAPDFAISDASGSDILGIEDMPQ 1666
 DB 1612 KYGSDLSGILNRAGIAPGINKHVEGSKPIFNALITPDLDAISEGSDILGIEDLPH 1671
 QY 1667 FRDDYFPPDGTSSGIVLHMGFATDNKPVTHGNNFGKSPYLPDQ 1713
 DB 1672 FGVNFMDEPVNEKANGY-----PEVANNNNNFAFNPRPFSPE 1708

RESULT 6
 P81137 PRELIMINARY; PRT: 1528 AA.

AC P81137; 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE INSCITICIDAL TOXIN RECEPTOR BT-R1 PRECURSOR.
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dityrysta; Sphingioidea; Sphingidae; Sphinginae; Manduca.
 NCBI_Taxid=7130;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=MIDGUT;
 RX MEDLINE=95197553; PubMed=7890666;
 RA VADIAJAGL-R-K., Medet E., Li L., Li T.H., Bulla L.A. Jr.;
 RT Cloning and expression of a receptor for an insecticidal toxin of *Bacillus thuringiensis*.
 RL J. Biol. Chem. 270:5490-5494(1995).
 CC -1 FUNCTION: BINDS TO THE CRYIA(B) TOXIN OF *BACILLUS THURINGIENSIS* SUSP. BERLINER, LEADING TO THE DEATH OF M. SEXTA.
 CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1 SIMILARITY: RELATED TO THE CADHERIN FAMILY OF CELL ADHESION MOLECULES. CONTAINS 11 CADHERIN-TYPE REPEATS.
 CC InterPro: IPR001525; C5_DNA_meth.
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin. 5.
 DR SMART: SM00112; CA: 10;
 DR PROSITE: PS00095; C5_MTASE_2; UNKNOWN_1.
 DR PROSITE: PS00232; CADHERIN_1; 1.
 DR PROSITE: PS50268; CADHERIN_2; 10.
 KW Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Cell adhesion.
 FT CHAIN 1 21
 FT DOMAIN 22 1528 INSECTICIDAL TOXIN RECEPTOR BT-R1.
 FT DOMAIN 72 1405 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1406 1428 11 X APPROXIMATE TANDM REPEATS.
 FT DOMAIN 1429 1528 POTENTIAL.
 FT REPEAT 72 176 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 177 289 CADHERIN 1.
 FT REPEAT 290 397 CADHERIN 2.
 FT REPEAT 398 500 CADHERIN 3.
 FT REPEAT 501 623 CADHERIN 4.
 FT REPEAT 624 757 CADHERIN 5.
 FT REPEAT 758 882 CADHERIN 6.
 FT REPEAT 883 1004 CADHERIN 7.
 FT REPEAT 1005 1121 CADHERIN 8.
 FT REPEAT 1122 1242 CADHERIN 9.
 FT REPEAT 1243 1353 CADHERIN 10.
 FT REPEAT 1353 1528 CADHERIN 11.
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 865 865 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 884 884 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 966 966 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1069 1069 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1528 AA; 171986 MW; CF678E01D700C91D CRC64;
 Query Match 52.0%; Score 4631; DB 5; Length 1528;
 Best Local Similarity 60.0%; Pred. No. 4e-287;
 Matches 915; Conservative 221; Mismatches 366; Indels 22; Gaps 13;

QY 1 MGVERFFPAVLVSLASALANDRCSTYITAIR-PETPELPIDYIGKWSQPLIPGT 59
 DB 1 MAVDVRIAAFLVETAPPAVLAQRCGYMTAIPRLPRDMLPVLENGQVWSQPLIPAPE 60
 QY REEVCMENF---LPDMQIVYMEEEIEGDVITIAKNTYQSNTPVLSI--MSGPAPAG 114
 DB 61 RDDLCDAVHVITANIGTQVITIMEDEIEITAIINNGPSTPFLFELPGSGSYL-LM 119
 QY 115 PERONEADGOMSLVITRODETATATQSVYSIQEGESQAVLALETIVINDNPPILQ 174
 DB 120 PYIRR--VDNGSASHHNAQHLELPMQMYMNVAVDGGSLVAGSLAVINDNAPIIQ 177
 QY 175 VVSACVYPHGEARLTDCVYQVSDRDEISTRMPTFRVDSRAADESIFYMNGEIDPSM 234
 DB 178 NPEPCRVPELGEGLTECTYQVSDADGRISTEFMTFRIDSVR-GDEETFYIERTNIPNM 236
 QY 235 FMKMTVGINSLNPEETQOLHFSVTAASDLSLNHTVMVQVENESPPRVATEFSVQ 294
 DB 237 MWLNMTIGVNTSLNFTSPHLHFSVALDSLPTHTVIMVQVANVSPRMLLEFAVQ 296
 QY 295 QPDEKTNQSFSLRAIDGIDGINRAINYTLIRDDADFFSLEYI---EDGAIHVTIEDRD 351
 DB 297 QREKSYQNFYRAIDGDEIMNPINRYRLITNEDEFFSIALPGKSAVFLVSPIDRD 356
 QY 352 KLERELNLTIVAYKSTQASPAETEAHIFTIVDNDQREPLHKEYSIDIMEETPLNLF 411
 DB 357 TLQREVFPLTIVAYKDEBAFSTSTNVVITVINDQREPLHKEYRIAMETPLTNLF 416
 QY 412 NEEFGRHDLGEMNQVETELDVPFGAASFYIAPGSGYORPTIMGTINMTMDYD 471
 DB 417 DKEFGHDKDLDQNAQYTRLSVDPGAAEFYTAPEVGRQRTITKTLNHSMDIYE 476
 QY 472 VIFQNIIRKAVKAVNNASHVGEALVYNLIMNDELPIFEESYSASAKETVAGCPVA 531
 DB 477 PFGQSTIRIVATDNNDRHVGVALVHIDLIMNDEQPIFEHAYQVTFDETEGEGFVA 536
 QY 532 TYLADRDIDVAVVSLMNAVDYLFIDSTGEIFYSMQADADYHQRNLPVQVADDTL 591
 DB 537 KAVAHDRDIDGVVEHLLGNANVFLTDKTDIRVASDENSEYHRESLFFQVARTDTL 596
 QY 592 GGPHTVITQVIELEDEVNNTPTLRPRSTPSVEENPEGEIETRETTADPTSAVL 651
 DB 597 GQ-PHRTATSQLIRLINDINNTPTLRPRSGPOVEENVPDAHVITQELRATDPTTADL 655
 QY 652 WEIWDSTWATKQGRNPTREYVOCIVETIYTPTEGNRSGAIGRLVVOEIDNVTIDE 711
 DB 656 REIIMWDTSEATKQGRQANPDEFRNCVEIETLEPEINNGLAIGRVAREIRHNVITIDE 715
 QY 712 EPEMLYLYRVADLTIVGDDYDEATFTTTIDANDNAPITFANGILTQMRARELAASGT 771
 DB 716 EPEVSLTVYRVADLTIVGDDYDESMITTTITIDMNDANVVEGTELEDFRFRSAGSL 775
 QY 772 LIGSVIATIDIDPLINOVYVITQPRNPEGLVKIDFTTGQLEVNANEAIDVDEPMRFL 831
 DB 776 VVGSRADIDIDPLINOVYVITFPREDIDKDLIMELPNKGSFRRHKRIDANTPPRRL 835
 QY 832 YVTIVIASDECSLENPTCEPDSNYFEVPGDIEIITDNKVPPELITKENTVYVWENA 891
 DB 836 YVTIVASDRCSLEPDADECPDPTIYETEGNITIHITDNNKVPQAEFTKFDIVVYENA 895
 QY 892 TSGDEVQVLSHRRDRELXHTVKTMTNVAVPRLDFFVEDLDGRLGVHPR---GDEK 948
 DB 896 THLDEVVTLIASDLDRDELXHTVKTMTNVAVPRLDFFVEDLDGRLGVHPR---GDEK 955
 QY 949 LDRDDEPTHTITFVNFIDNFSDGGRNQDEVEIFVVLVDVNDNAPEMPLPDELRFQVS 1008

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:34:02 ; Search time 7575.83 Seconds

(without alignments)
11972.483 Million cell updates/sec

Title: US-09-715-909-1
Perfect score: 5498
Sequence: 1 cataatacaataagaaga.....aaaaaaaaaaactcgag 5498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
1 number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hig:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hcgo_hum:*
31: em_hcgo_inv:*
32: em_hcgo_rod:*
33: em_hcgo_hum:*
34: em_hcgo_inv:*
35: em_hcgo_rod:*
36: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5498	100	5498	6	AX147201	AX147201 Sequence
2	20178	36.7	3582	3	AF319973	AF319973 Manduca s
3	1982	54.69	5469	3	AB026260	AB026260 Bombyx mo
4	1975.2	35.9	5172	3	AB041508	AB041508 Bombyx mo
5	1964	35.7	5172	3	AB041510	AB041510 Bombyx mo
6	1959.2	35.6	5172	3	AB041509	AB041509 Bombyx mo
7	1954.6	35.6	5577	6	AR096316	AR096316 Sequence
8	1954.6	35.6	5577	6	I77078	I77078 Sequence 1
9	1946	35.4	5355	3	AF367362	AF367362 Heliothis
10	1826.6	33.2	5527	6	AX147203	AX147203 Sequence
11	1736	31.6	5527	6	AX150183	AX150183 Sequence
12	1547.4	28.1	5604	6	AX147207	AX147207 Sequence
13	1465.8	26.7	5592	6	AX147205	AX147205 Sequence
14	1363.2	24.8	7799	3	AE367363	AE367363 Heliothis
15	61.8	1.1	11419	1	AE004760	AE004760 Pseudomon
16	58.4	1.1	52863	2	AC017834	AC017834 Drosophill
17	58.4	1.1	174500	3	AC009202	AC009202 Drosophill
18	58.4	1.1	251154	3	AE003656	AE003656 Drosophill
19	58.4	1.1	294942	3	AC011662	AC011662 Drosophill
20	57.2	1.0	10087	1	AE005045	AE005045 Halobacte
21	56	1.0	35856	1	SC59	AL138851 Streptomy
22	55.2	1.0	37931	1	SCD10	AL135988 Streptomy
23	55	1.0	1319	1	STMIMP	D00670 Streptomyce
24	55	1.0	1899	1	STMIMP	M89476 Streptomyce
25	55	1.0	2007	1	AB042262	AB042262 Corynebac
26	55	1.0	2045	1	STMENDOPRO	M81703 Streptomyce
27	55	1.0	2261	6	I20933	I20933 Sequence 1
28	55	1.0	41807	1	SC6D11	AL158061 Streptomy
29	54.6	1.0	196490	2	AC005507	AC005507 Plasmodiu
30	54.2	1.0	3051	1	ABU20583	U20583 Azospirillu
31	54	1.0	13040	1	AE005769	AE005769 Caulobact
32	54	1.0	152973	2	AP003334	AP003334 Oryza sat
33	53.4	1.0	24800	1	SC9B1	AL049727 Streptomy
34	53	1.0	22775	1	SC5G8	AL133872 Streptomy
35	53	1.0	36144	1	SC9C5	AL137523 Streptomy
36	53	1.0	113193	1	AF357202	AF357202 Streptomy
37	52.8	1.0	10982	1	AE005834	AE005834 Caulobact
38	52.6	1.0	123580	1	AF263912	AF263912 Streptomy
39	52.4	1.0	37123	1	SC5F7	AE06872 Streptomy
40	52.2	0.9	11623	1	AE002066	AE002066 Deinococc
41	52.2	0.9	12895	1	AE008084	AE008084 Agrobacte
42	52.2	0.9	14860	1	AE007093	AE007093 Mycobacte
43	52.2	0.9	63033	1	MTV008	AL021246 Mycobacte
44	52.2	0.9	199551	2	AC006281	AC006281 Plasmodiu
45	52	0.9	1141	8	HVU26545	U26545 Hordeum vul

ALIGNMENTS

RESULT 1
LOCUS AX147201 5498 bp DNA
DEFINITION Sequence 1 from Patent WO0136639.
ACCESSION AX147201
VERSION AX147201.1 GI:14346372
KEYWORDS
SOURCE
ORGANISM European corn borer.
Ostrinia nubilalis
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Pyraloidea; Pyralidae; Pyraustinae; Ostrinia.
REFERENCE
AUTHORS Clannagan, R.D., Maents, J.P. and Meyer, T.J.
TITLE Bt toxin receptors from lepidopteran insects and methods of use
JOURNAL Patent: WO 0136639-A 1 25-MAY-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1..5498
/organism="Ostrinia nubilalis"

Applied

Db 1501 CCCCGGGAGCGCTACAGAGAGACCTTCATCATGGGACACATTAACACACCATG 1560
QY 1561 tggattacgaagatgtcatcttttcagaacatcatatcaatgaagtcgaagtcgacatga 1620
Db 1561 tggattacgaagatgtcatcttttcagaacatcatatgaagtcgacatga 1620
QY 1621 acaagcgaagcagcttgagcagagcgtgtgtacgtgaacgtgacatgaagcagc 1680
Db 1621 ACAAGCCTAGCAGCTGGGCGAGGCGCTGTACTGTGACCTGATCAACTGATGGAAGCAG 1680
QY 1681 aactcccaacttcgagagagcagctactccgcgtcgtttaaggagacgcgtgcgcgc 1740
Db 1681 AACTTCCCATCTTTCAGAGAGAGACACTACTCCGCGCTTTAAGAGAGACGCTGGGCGCG 1740
QY 1741 gcttcccggtgacagcgtgtgtcgcgccttcgacagagacatgaagcgtgaagtcatt 1800
Db 1741 GCTTCCCGGTGACAGGCTGCTGCGCTTCGACAGAGACATCAGACAGCTGATGTCATT 1800
QY 1801 catgtatggcagcagctgtgtactactgttcataatgaatcaacgaggagatcttcg 1860
Db 1801 CATGTATGGGCAACGCTGTGTGACTACTCTTTCATAGATGAATCAACGGAGAGATCTTTCG 1860
QY 1861 tgaagatggagcagctgcttcgactacacacagacagacatctatctgttcagctgcgc 1920
Db 1861 TGAGCATGGAGCATGCTTCGACCTACACACGACAGACACTCTATTTGTTACAGTGGCGG 1920
QY 1921 ctgaagcattcttggtgagcagcgcgcacacacacagtgacacacagctgtgtatagaac 1980
Db 1921 CTGACGATACCTTTGGGCGAGCGGCCACACACACAGTGCACACGCTGATGATGAAC 1980
QY 1981 tggagatgtcaacaacactcctccacacactcgccttgcgcgcgttcgactcgaagcgttcg 2040
Db 1981 TGAGAGATGTCAACAACACTCTCTCCACCTTACGCTTCCGCTTCGACTCCAAAGCCTTCG 2040
QY 2041 aggagacgcttcgcgaagatagacagatalccggggaatcaactgtctacacacccgcgaca 2100
Db 2041 AGGAAAGCTTCCCAAGATACGAGATATCCGGGAATCACTCTACCCGACCCGAGACA 2100
QY 2101 cgaagcctactctgtgtgagatcgaactggaactcgaactcgggcgcacacacagcgggaca 2160
Db 2101 CGAGGCTTACCTGTGTGCTCGAGTGCAGTGGACTCCACTGGGCGCACCAAGCAGGGACA 2160
QY 2161 gagagaccaaaccctactaatacgtcgcggtgtatagttatcgaacagatalaccacacgc 2220
Db 2161 GAGAAACCAACCCCTACTGAATACGTGGGTGTATAGTTATCGAAACGATATACCCACCG 2220
QY 2221 agggcaacggggttcgcgcacatcgggcgcctcgtgtgcgaagagatccgcggacaacgttca 2280
Db 2221 AGGGCAACCGGGGTTCCGCAATCGGGCGCCTGTGTGTCGAAGAGATCCGGGACAACTGCA 2280
QY 2281 ccatcgaacttcgagaatctcgagatgctttactactacacgctcgcgttgaggacctcaaca 2340
Db 2281 CCATCGACTTCGAGAAATTCGAGATGCTTTTACCTACACGCTCGCGTGAAGGACCTCAACA 2340
QY 2341 ctgtcatcggagatgaactacgaatgagcgaagcttcaagatcaacaataatcgacatgaagc 2400
Db 2341 CTGTCTATCGGAATACATCAATGAGGCGAGCTTCAAGATCAATTAATCGACATGAAGC 2400
QY 2401 acaacgcgcgcagcttcgcgaagcgcgcgtgcgcgaagcagatgcgcgtgcgcgcgcgcgc 2460
Db 2401 ACAACGCGCGCGATCTTCGGAACGCGACGCTGACGACAGATCGGCGTGGCGGAGCTGG 2460
QY 2461 cggcgaagcgcgcagcttcgcgcgtgcgcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2520
Db 2461 CGGCGAGCGGACGCTCTATCGGCTCCGCTCGGCTCGACACGACATCGACGCGCGCTCTCA 2520
QY 2521 accaagctgcctacactatacaacactagaacaacacactcccgaggatgaatgaagtcg 2580
Db 2521 ACCAAGTGGGTACACTATACAACTAGAAACAACACTCCGAGGAGTTATGAGAAATGG 2580
QY 2581 acttcaacactggtcnaattgagtgatgcgaagcagcgtatgcgaagcgaacccct 2640
Db 2581 ACTTCAACAAGTGTCAAAATGAGTGGATGCAACGAGGCGATCGATCGACGCAACCT 2640

QY 2641 ggcgcttctactgttactacacccgtcatcgtctagcgaagctgtccctcgaaacgcga 2700
Db 2641 GCGGCTTCTACTTGTACTACACCGTATCGTATGCAAGAGTCTCCGTGAAAAACCGCA 2700
QY 2701 cggaaatgtccctccagatcccaactactcgaagttccaaagcgaatcgaatagaataca 2760
Db 2701 CGGAATGTCTCCAGATTTCCAACTACTTGGAAAGTTCCAGCGCATATCGAAATAGAATCA 2760
QY 2761 tcgacacaaacaagaatgtccttgagcgcgtcacttgagaagttcaacacgcgtgtacg 2820
Db 2761 TCGACAAACAACAAGAGTCCGTGAGCCCTCATGAGAAAGTTCAACAGACAGGCTGTACG 2820
QY 2821 tctggagagaatgcacacgc 2880
Db 2821 TCTGGAGAGATCCACAGACGCGCGACGAGAGTGTCTCAGCTGTATCTCCACGACCTGTACA 2880
QY 2881 ggaacgaatgtgacacacagcagatagacatacgaatgaacttcgcgttgaaacccgcagtcg 2940
Db 2881 GAGAGAGTGTGACACACAGGATACGATACGATGACTTTCCGCTGGAACCCCGACTGC 2940
QY 2941 gggattcttcgaagtgagaccttgacacactgctgcgccttgaggttgacatcccgaggagc 3000
Db 2941 GGGATTTCTTCAGAGTGGACCTGGAACACTGCTGCGCTTGAAGTGCATTCACCGGGGAGC 3000
QY 3001 aaaaattgagcgcgcgtatgggagatgagcctacacatactatcttbaattcatcgata 3060
Db 3001 AAAAAATTGAGCCGATGGGGATGAGCCCTACACATATATCTTTGTAATTCATCTGATA 3060
QY 3061 acttcttcttgatgtgtacagcgaagagagaacacagagcgaagtgaatattgttcgttc 3120
Db 3061 ACTTCTTTCTGATGTGTGACGCTAGAGAGAAACAGAGCAAGTTGAATATTTGCTGTTTC 3120
QY 3121 tatgtatgtgaacgaacacgcctctgagatgcaltgctcgtgaatgcagctcggtttatg 3180
Db 3121 TATGTGATGTGAACGACAACGCTCTGAGATGCATTGCTGATGAACTCCGCTTGTATG 3180
QY 3181 ttctcgaagagcgaattctgtgtgtcgtgtgtatccagaagaatctacgcacacgcgcgcgc 3240
Db 3181 TTTCCGAAGAGACCAATGTGTGTGTCGTTACTCCAGAAATCTACGACCGGACAGGG 3240
QY 3241 atgaacacagacacgcgcacacactcgcgtgtcgcgttgaagatcccttgacacacacgc 3300
Db 3241 ATGAACCAAGACAGGACCAACTCGCTGTGCTGATGAGATCTTGAGACCTGACAGATCACCG 3300
QY 3301 accgagacatcgaagtgccgcgactctctccacatgactcgaatgaaacaaactgggg 3360
Db 3301 ACCGAGACATCGAGGTGCGCGGATCTTTCACCATGATCTCGATTGAAAAACAACCTGGGG 3360
QY 3361 aacttgaacgcgcataatgacttgagggggtatgtggcgaacttaagaataatcatatgag 3420
Db 3361 AACTTGAACCGCTATGGAAGCTTGAAGGGGTATTTGGGCACTTACCAAAATATTCATTTGAG 3420
QY 3421 ccttcgacacagcgtcaccgc 3480
Db 3421 CCTTGACACCGGCTTACCCGACGAGAGTCCAAAGAACGTCACCTGTGATCCGCC 3480
QY 3481 cctacaacttccacacccctgtgtcgtgttccgcgaacccgcgaactcgtlcatctggctt 3540
Db 3481 CCTACAACCTTCCACACCCCTGTGCTGCTGCCGCAACCCGACACTCCGATTTGCGCTTT 3540
QY 3541 ctgaggagcgcgcgaacaaagagcgcgcgttcttggcgaacgcgttcgcaagagttccctggagc 3600
Db 3541 CTAGGAGCGCGCAACAAAGCGGCGTGTGGCGACGCTGCCAAACGATTCCTGGAGC 3600
QY 3601 cgaatcagccacacgc 3660
Db 3601 CGATTTAGCCACGAGCAGAGAGCGGCTCCACGCGGAGGCTACGTTCCAGCTCCAGG 3660
QY 3661 gaaatgagagcgcgttcgaacttgaataactgaagtggagcagagagaatagcgc 3720
Db 3661 GAAATGAGAGGCGCTTCACTTGTGATTAACGAGTGGAGACGAGGAAATATAGCG 3720

QY 3721 ggcagctatatactacgcaagctttcccaagacaatcagaacaatcagaatcc 3780
 |||||
 Db 3721 GGCAGCTTATATATACGCGAGCTTTCCAGAGCAATACAGCAATACAGATCC 3780
 |||||
 QY 3781 gggcgaagagagcgagcgagcgccgctttgagcgaagtaagtttcggtg 3840
 |||||
 Db 3781 GGGCGACGAGGCGGAGAGCGCCCGCTTTGGACGAGTCACTGTTTCGGTGG 3840
 |||||
 QY 3841 tcttcgtaccccaacagggcgaccagtggttcacgaaatgacgtactctcctct 3900
 |||||
 Db 3841 TCTTCTGATACCAACAGGCGGACCCAGTGTTCAGCGAAATGACGCTACTCTCCCTTCT 3900
 |||||
 QY 3901 tcaagagtgaaagagcgccctcgtagagtttgaagctgacgcaagcagaacctaa 3960
 |||||
 Db 3901 TCGAGGTTGAAGAGCGCTCGTAGAGTTTGGAGCTGCCCAAGCAAGCACTTTAAA 3960
 |||||
 QY 3961 accacccctcgaagagagctgcaagatatctactacagttatttgaagcgaacaa 4020
 |||||
 Db 3961 ACCACCTCTGGAAGATACGCAAGATATCTACTACAGGTTTATGAGCGCAACAG 4020
 |||||
 QY 4021 agggctcttctgaactgacagtcacgaacgtcatctcccttgcgaggaagtgaac 4080
 |||||
 Db 4021 AGGGCTCTTTGCTACTGAGACCAAGTCAAGCAAGTATCTCCCTTCCGAGAGATTGGAC 4080
 |||||
 QY 4081 gcgagtgagcagctctacacgctcgcacatcgcgagcaactcgcccgacgacatg 4140
 |||||
 Db 4081 GCGAGTGCGCCAGCTTACACGCTGCACATCGCGGAGCAACTCGCCCAAGCCACTG 4140
 |||||
 QY 4141 ggtatccctctgacagacttccctctcgttgcacaggtcaatgtaagaaagcgaaccc 4200
 |||||
 Db 4141 GGATCCCTCTCAGACTTCCCTCTGTTGACGGTAAATGTAAGAAAGAAAGCCGCG 4200
 |||||
 QY 4201 gcccaattctcgaagcagacatttacacagcgagcgatctcagcttgacagacattg 4260
 |||||
 Db 4201 GCCCAATTTTGGAGAGGACCTTTACACAGCGGGGCAATTCGAGCTTGACAGCATTTGACC 4260
 |||||
 QY 4261 gggagattgcttaactgctcagcgagcgacacagaagaagacacacacacacacac 4320
 |||||
 Db 4261 GGGGATTTGCTTACTGTCAGGGGAGCCACACAGAAAGAACACCACTACACACCATAG 4320
 |||||
 QY 4321 accgtcgagatcagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 4380
 |||||
 Db 4321 ACCGTGAGATGAGCTGAGTGAACAGACCTTAGAAGCCGTGCGGACTCGGCTTTCGGGC 4380
 |||||
 QY 4381 tgcatacgacacacgcgctgtcttgcgtcaatgacagcagcagcagcagcagcagc 4440
 |||||
 Db 4381 TGCATGCGACACCGCGGCTGCTTTCGCTCAATATGACACCCACCGCTTCATGACGGCA 4440
 |||||
 QY 4441 tgttcgagtgcaagcgtacgctacgatalacagcttctgcaatcgacacagccgltg 4500
 |||||
 Db 4441 TGTTCGAGTTGACGCTACCTTACGATGATACAGCTTTCGCAATGACACAGCCGTTGCA 4500
 |||||
 QY 4501 aagctactctatctcactcgcaaaacgcgagcttcatittgataacacacttga 4560
 |||||
 Db 4501 AAGTCTACTCTATCTCATCGCAAAACCGCGTGAACCTTATTTTCATTAACCAACTTGA 4560
 |||||
 QY 4561 ccgttgagcagaacgaatctcagcgccagctcagcagcggttcaacatgaagt 4620
 |||||
 Db 4561 CCGTTGACCAAGCAAAATTTTCATAGCGGCCACTTTCAGACCCGGGTTTCATGACAGCT 4620
 |||||
 QY 4621 gcaacatcgacacagtggtgctgcgttcagcagcagcagcggtgagcagaacga 4680
 |||||
 Db 4621 GCAACATGACACAGGTGCTGCGTTTCAGCGACAGCGGCGTGGCGCAAGACGACACCA 4680
 |||||
 QY 4681 ccgaggtgctgcgagcttcatctcgggagacagctccgttgagggcacaagaggtcg 4740
 |||||
 Db 4681 CCGAGGTGCGGCGACCTTCCGCGAACAGTCCCGTGAAGCAAAAGGTCGAGG 4740
 |||||
 QY 4741 ccgttcgagcagaacagctgtcgtgcgcacacatcagctgtgttgagcaacacac 4800
 |||||
 Db 4741 CCGTTCGAGCAGACAGCTGCTGCTGCGCACCATCTACAGCTGATGCTGAGACCAACAGC 4800
 |||||
 QY 4801 tgggtctgcaagaccttggtagcgggtgcaactcgacgctaggcgaagagatcaatg 4860
 |||||

Db 4801 TGGTGTCTGCAAGACCTGCTGACGGGTGACACTCCGACCTTAGGGGAGGATGAATGCA 4860
 |||||
 QY 4861 tgcgcgtctacgaactagccgcgctctcgcgtgtgctgagcttctcctcgtactgc 4920
 |||||
 Db 4861 TCGCGCTTACGCACTAGCCGCGCTCTCGCTGTGTGCTAGGCTTCTCTCGCTCGACTGC 4920
 |||||
 QY 4921 tcttcgactgtctctgtgagcaagaagcactgaacagcgagctgaagcacttccatga 4980
 |||||
 Db 4921 TTTTCGCAATTTGCTGTGAGCAAGAGCACTGAACCGGCGACCTGCAACACTTTCATGA 4980
 |||||
 QY 4981 cgaagtacgctgtgtagctcccgagcggtgaacccgagcgcgctgagcgagcacaaca 5040
 |||||
 Db 4981 CGAAGTACGGCTCGGTGAGCTCCGGCTGAACCGGCGGCTGGCGCGGACCAACA 5040
 |||||
 QY 5041 agcaagcgtcgaagggctcccaacccatgttgaaagagcagcagcagcagcagcagc 5100
 |||||
 Db 5041 AGCAGCGCTGAGAGGCTCCCAACCCCATGTGGAAGAGAGCCATCGCGCGGCACTTG 5100
 |||||
 QY 5101 acgcacatcagtcagcgagtgagctcgcacacatcagcagcagcagcagcagcagc 5160
 |||||
 Db 5101 ACGCATCAGTACGAGCGAGTGGCGACTCCGACTGATCGCATCGAGCAATGCGCAAT 5160
 |||||
 QY 5161 tccgcgaagcacttctcccgcccgcgacacagactcaagcagcgagctgtctgaca 5220
 |||||
 Db 5161 TCCGCGAGACTACTTCCCGCCCGGCGACAGACTCAAGCAAGCGGCAATGCTTGCAHA 5220
 |||||
 QY 5221 tggcggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5280
 |||||
 Db 5221 TGGCGAAGCCAGCAGCAAAAGCCGAGACAGCATGGCAACAACTTGGGTTCAAGT 5280
 |||||
 QY 5281 ccaaccgctacctcgcacagcgcccaagaatcctcagggatataactgttccagg 5340
 |||||
 Db 5281 CCACCCGCTACTGCTGCACAGCGCACCCAAAGTACTCCAGGATTAACCTGTCAGGG 5340
 |||||
 QY 5341 tgcctacgcgcgagagtgcgacacagcggttatacagcagcagcagcagcagcagc 5400
 |||||
 Db 5341 TGCTTACGCGCGGCAAGTGGCAGACCGCTTATCATTCGGGAACATTAGCATGAAGT 5400
 |||||
 QY 5401 acctatgacatattgtaaatgtatcacatlatcttatacaaatatatttat 5460
 |||||
 Db 5401 ACCTATGATACATATTGTAATGTAAATATCTATTATTAACAATATATTATTAT 5460
 |||||
 QY 5461 attgctaaaaaataaaaaaataaaaaaactcgag 5498
 |||||
 Db 5461 ATTTCCTAAAAAATAAAAAAATAAAAAAATCTGAG 5498
 |||||
 RESULT 2
 AF319973
 LOCUS AF319973 5582 bp mRNA INV 05-DEC-2000
 DEFINITION Manduca sexta cadherin-related protein receptor Br-R1 mRNA,
 complete cds.
 ACCESSION AF319973
 VERSION AF319973.1 GI:11545673
 KEYWORDS
 SOURCE
 ORGANISM
 tobacco hornworm.
 Manduca sexta
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 Ditrysia; Spingulodea; Spingulidae; Spingulinae; Manduca.
 REFERENCE
 1 (bases 1 to 5582)
 Dorsch,J.A., Maaty,W.S.A., Griko,N.B., Candas,M. and Bulla,L.A. Jr.
 A cadherin-related Protein Receptor, Br-R1, in the Midgut
 Epithelium of Manduca sexta Mediates Toxicity for Bacillus
 thuringiensis CryIIa Toxins
 JOURNAL
 Unpublished
 2 (bases 1 to 5582)
 Dorsch,J.A., Maaty,W.S.A., Griko,N.B., Candas,M. and Bulla,L.A. Jr.
 Direct Submission
 TITLE
 Submitted (07-NOV-2000) Center for Applied Biology, University of
 JOURNAL
 Texas at Dallas, Richardson, TX 75083, USA
 FEATURES
 Location/Qualifiers

source 1..5582
/organism="Manduca sexta"
/db_xref="taxon:7130"
CDS 197..5530
/function="receptor for Bacillus thuringiensis CryIA
toxins"
/note="midgut epithelium cadherin-related protein
receptor"
/codon_start=1
/product="cadherin-related protein receptor BT-R1"
/protein_id="AAG37912.1"
/db_xref="GI:11545674"
/translation="MAVDRIAAFLVFIAPAVLAOERCGYMTAIPRLPDPNPLVN
FEGTWSORPLPAPERDLCMDAYHTVATNGVIVIMDEIEIEITLITLNTGSLV
TPETLPLSSSYNLMVIRVNDNGEHLITTOQHLEIDPMQOYMNVRVDSGLV
AGVSLAIVNDNAPITONFECRVELGEGLTCTYQVSDADRISTEETMRTIDS
VRGDEITYIERINIPNQMWMNMTIGVNTSLNFTSLHLEISVALDSLNTHTVM
MVQAVNYSRPRMLEFAVOOEKSYONFTVRAIDEDTEINMEINRYLITNEEDF
FSIEALPGSGKGAFLVSPIDRLQREPLTVAAYKDEASTNNVLIYDIN
DORPEPIKEXRLAIMEETPLTNFDEKGDHDKQAOYTVARLESVDGAAEAF
YLAPEVGRQTFINGTLNHSMLDYEVEPESITIRVATPDNNDRHGVGALVHTDI
NMNDEQPIFEHNAVOTVTFDEBEGEFPAAKVAHDROIDGYEHTLGNVNTLIDK
LNGDIRVANDSENFHRESELVQVRAIDTLEPEHTASQVLRINDINNTPLRL
PRGSPQVEENVDPDHVITQELRADPDPTADLRREINDSEFATKQGNADPERNC
VEIETIPEEINNRGLAIGRVVAREIRHNVITIDEEFELSLTVRVDINTYGDYDE
SMLTITIDMNDNPVMEGTLEONFRVREMSAGLVGVSRADDIDGPELNOVYTI
FPREDPKDLIMIDFLGOISVNTSGAIDADTPPEHLYTVASDRSTEDPACCP
DPTVETEGNITHTIDTNKVPQAEETKEPTVITYENATHLEIVTLIASDRODE
IHTVSVINAVNRLNMFVSVNREGLVYVDETQSGEVLDRDEPHRIEFLN
IDNMEGEGNRNODNEVLVILDVNDNAEELPPSELMTISENLKOGVRLPHIT
APDEPDENSRVYELINLSTEDIDYELFVLIQANTGELETMDKVGWGY
AIIHRAFDHGIROMSMNETYELIHPEVYVPEFVINDVIRLARERAYINGLAT
VNGELIERISATDPDGLHAGVYTFQVYVDESGRFOVYVNGENGLSRILQANDEI
REPRTIRATPOGDPGLSDMTFRVYVTFQGPFRASSEHVAFLFKESAGBESH
OLPAAODIKNLCEDDCHSTYIRIIDGSEGHGLDPVNRMLFKELIRQDSHHTL
QVAASNSPDGILPLPASLITVTVYREADPPRVFRELTYLGIISADIGELRLHA
TOSBSAITYAIDYDTMVDVPSLEAVROSALVNAQVTLNIOPTMGLKFEV
TATPDAGODRDTVYVYVSSONRYVFEVNTLOOEVRNDEIDFEGANVINDID
OVVAPNDPVGTVALHESNOMRGHFRNDVPVLADEIRIDVILDSIQTTNARSL
VLQDLITNSSPDSAPDSLTYYVLAISAVIGMCLVLLFTITRLANRLRLSLM
TKYSGLSBGLNRAGIIAAGTNKHTVEGSPNFENEXIKPDIADLASEGNSDDLIGIED
LPHFGNFMDEVEVNEKANGPEVANHNHNFNPTPSPERFVNQFRKI"

BASE COUNT 1478 a 1499 c 1405 g 1200 t
ORIGIN

Query Match 36.7% Score 2017.8; DB 3; Length 5582;
Local Similarity 64.1% Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 1767; Indels 57; Gaps 12;

QY 152 atcagcaaaatgggggttgagaggttcttcccaagcagtcgactggtctcttgcctc 211
DB 187 AATAATGTGAGATGGCAGTGGACGTCGGAATCGCTCCCTCCGCTGGTGTATATAGCGCC 246
QY 212 tgcgcactagcaacaacagatgttcgtaactatcgcaataccaagacggagactc 271
DB 247 TGCATTTTACGTCAGAGAGATGTGGATATGACCGCATCCCAAGGCTACACAGCAC 306
QY 272 ggaac--tgcgcactatgtattcgaaggaataatcattgagtgagtgacagcctctaacc 328
DB 307 GGAATATTGTCGACAGTAAATTTTGAAGGCCAGACATGAGACTGAGAGCCCTCCTCC 366
QY 329 cggccgagcccgagaggaagtatgtatgagagaact-----tctaacggatacaat 379
DB 367 CGCCCGGAGCGGATGACCTGTGATGACGCCCTAACACGTGATTAAGCCAAACCTCGG 426
QY 380 gattcaggtcatatatacgaaggaagaatcgaagagacgcatcatctcgaagctaa 439
DB 427 CACGAGAGGTCACTTCATGATGATGAAGAGATAGAAACGAATACCATTCGCTATCTTAA 486
QY 440 ctataaaggttcaacaacgctgctgctgattatgtcaggccagccagacagct 499
DB 487 TTATAACGACCATCACTCGCTTCATTGAACCTGTCATTTCGCGGTGTGTAACATCT 546

QY 500 gggccctgagtttcgacacgaatgaagcagcgcccaatgagcctgttattacgcaag 559
DB 547 GGTGATGCCGGTCAATGAGAGATTGACAACGGGGAGTGGCATCTCATCATCGCAAG 606
QY 560 acaagactacgagacagacaacatcagagctatggttctcaatccaatgagagtg 619
DB 607 ACAGCATTTACAGAGTTCCCGGCATGAGCAGATGACATGTTTCATGTCCGGTGACGCCA 666
QY 620 atcaagagccgactggtggtgcgtgagagatagcaatcagacgacaatccgccaatc 679
DB 667 GTCGCTGGGAGCGGCTCTCTCTGCTATCGTCAACATAGATGACAACGCCCCATCAT 726
QY 680 gaaagtgtgcaacgctgctgtaattccagaacatgagcgagctagactgaccgactg 739
DB 727 ACAAAATTGAGACCTTGCGGGGTCTCTCACTGGCGCCAGCCAGGTTTACGAAGATGCAC 786
QY 740 gtaccaagtgtcagacgcyagcgtgtaaatcagaccgacctcaatgacgtccgtgca 799
DB 787 ATACCAAGATATGAGACGGCGAGCGGATGACACAGAGTTCAATGACGTTGAGATGCA 846
QY 800 cagcagcagggctcgaatgaaagcaatctctacatggttggagaatcagaccceagca 859
DB 847 CAGCGTTG---TGGCGACGAGAGACCTTCTACATCGAACGAGCAATATCCCAACCA 903
QY 860 ctggttcaatatgaaagtatgctggtggaatcgaatccgcttgaaactgagacaactca 919
DB 904 ATGAGATGTCGCTTAATATGACATAGCGCTTAATCTCTCACTTCACTTGTCAACAGTCC 963
QY 920 gctcatatattagcgtcacaagcttctgactcgtcactaccgaacaacacagcgtcacat 979
DB 964 GGTGATATATTACAGGTGTACAGCCCTGACGCTCCCGGAACCCACGCGTACTTAT 1023
QY 980 gatggtcgaagtggagacgtagagctcgcgcctccgctggtggtgagatcttctcagt 1039
DB 1024 GATGGTGCAGTGGGGAATGTGAACAGCCGCTGCGCGCTGAGATCTTGCTGCT 1083
QY 1040 ggaagagtttgaaggaagacatcaagactcgcctcccgcgagatagcggagacac 1099
DB 1084 CCAACAGTTTGAAGAGAAATCTTACCAAACTTCCACAGTAGAGGCGATGAGAGGACAC 1143
QY 1100 ggaatcaataagggcatcaactataccctcaagagatgtagcgtgtagaactcttcc 1159
DB 1144 TGAGATTCATATATGCTATATCACTATACAGGCTGATCAACATAGAGAAACATTTCTGAG 1203
QY 1160 cctggaagtgatg-----aagacggagctatctcgcagtgaactgtagacgacg 1210
DB 1204 CATTGAGGCGCTGCTGCTGAGAAAGCGGGGCTGTATCTCTGTCGCCAATTGACCG 1263
QY 1211 cgaacagcttgaagaagagcttccaactcacaatcggttgccttaaatcaactgagacg 1270
DB 1264 CGACACACTGCAACAGAGGTGTTCCACTTACGATCGTGGCTTCAAAATATGATGAGGA 1323
QY 1271 tagcttgaacagagggccacattctcaatcgtaacagcgtcaatgatacagacg 1330
DB 1324 GGCTTTCACATATCAACAAAGTGGTATCATTTGTGACAGACATTAAGACCAAAAGAC 1383
QY 1331 cgaacgctgtcaataagaatagatgatatcaatggaagaaactccaatgactctaa 1390
DB 1384 TGAACCTTATACACAGAAATATCGACTGCAATGAGGAGAGCGCCCTGACCTCA 1443
QY 1391 cttaataagaagaatttgatllcatcagatcagagatttgggtgaaaaacgctcaataacagt 1450
DB 1444 CTTCGATTAAGAAATTCGATTTTCATGATTAAGATTGAGTCAAAAACGCTCATATACAGGT 1503
QY 1451 ggaacttgaagagcgttcccgccagggcggtccgcatctcactacgcgcggggag 1510
DB 1504 GCGTTTAAGAGAGCGTGGACCTCCAGGGGCTGCTGAAGCATTTCTACATAGCGCCGTAAGT 1563
QY 1511 cggctacagagagcagaccttcaatcattggtgacacataaacaacacatgctgattacga 1570
DB 1564 CGGTACACGAGACAGACCTTCATCATGAGGACACCTCAATCACTCCATGCGGATTAAGGA 1623

Oy	1571	agaagtcatttttcgaaacatcatlaaagtcgaagcagtagacatgaaacacgctag	1630
Oy	1571	agaagtcatttttcgaaacatcatlaaagtcgaagcagtagacatgaaacacgctag	1630
Dp	1624	AGGCCAGAACTTTTTCACAGATTATACGATTTGGGTGGTGTACCGACGCAACAACGACACGAG	1683
Oy	1631	caacgtgagagagcgctggtgtaagctgtaaccctgtacacgtgaacgcgaacgcgaactcccat	1690
Dp	1684	GCACGTGGGGCTGGCTGTGGTTCACATTGACCTCATCAATTGGAAACGATGAGCCGAT	1743
Oy	1691	cttcgagagagacagctactccgcgctcgttlaaagagacgctcgcgcgctccgcgt	1750
Dp	1744	CTTGGAACACGCGGTGGCAGACGGTCACCTTGACGAGACTGAAGCGAGGGGTCTTCGT	1803
Oy	1751	gagcaagctgctgcctccctgcgaagagacatgcgaacgtagtagtgcattatgtatggg	1810
Dp	1804	CGCCAAAGCGGTGTGCACACAGACAGAACATCGGGGATGTGTGTGAGCATTTATTATGGG	1863
Oy	1811	caacgcgtgtgactacgtcttcaatagatgaatcaacgcgagagatcttcgtgcgacatgga	1870
Dp	1864	TAAOCGTGTTAACTTCTGTGACCATGACAAATCTCACGGCGACATCGCGCTCAAGCTAA	1923
Oy	1871	cgatgctctgcactaccacgcgaacgaacactatctgttcaagctgcgcgtgcgaatc	1930
Dp	1924	CGACTCTTCAACTACCATCAGACGAGAAAGTAAATTATTGTGACAGGTGCGAGCTACGACAC	1983
Oy	1931	tttggcgacgcgcccacacacacacagctgacccacacgctgtagatagaactcgagagatgt	1990
Dp	1984	GCTGGGGGA---ACCTTTCACACCGCGCAGCGTCACAGCTGGTCATACACATTAATGACAT	2040
Oy	1991	caacaaacaccccccacacctaagctgtgcgcgccttcgcactcaacgcgtcgagagagacgt	2050
Dp	2041	CAACAAACGCCACCCACCTTACGGCTGCTCGAGGACATCCCAAGTGGAGGAAAGCT	2100
Oy	2051	tcgccgaagatacagagatalcccgaggaaatcaactgtcactacgcacccgacacagcgctca	2110
Dp	2101	GCTGATAGGCCACGTCATACCCAGGAGTTACGGCGCACCGACCCGACACACGCGCGA	2160
Oy	2111	cctgtgtctcgagatcgaactggtggaactcaactcgccgacacgaagcagagcagaccaa	2170
Dp	2161	TCTCCGCTTTCAGATTAACCTGAGGACACCTCTTTCGCCACAAAGCAGGCCCGCAGGGCTAA	2220
Oy	2171	ccctacacgaatgcgtgggtgtagtattgcgaacgatatccccacgcgagggcgaacg	2230
Dp	2221	CCCCGACGAGTTTAAAGAAATTGCGTGGAAATTCAGACCATCTTCCCCGAGATTAACCAACG	2280
Oy	2231	gggttcgcgcacatcgggcgacctcgtgtgtgcgaagatgccgggacaaacgctcacactgcact	2290
Dp	2281	GGGACGTGCTATCGGGCGGGCTGTACGCGCGCAAAATTCAGACCAACGTCGACATAGACTA	2340
Oy	2291	cgaggaatcgcgaatgcttactctcaacgcgtccgcgtgtagaggaacctcaacactgtcaccg	2350
Dp	2341	CGAGGAACTTTGAAGTCTCTCTCTCCACAACTGAGGGGTGGCTGACCTTAAACCGCTTACGG	2400
Oy	2351	agatgactagatgaagagcgacgtctcaacgtatcaataatcgcacatgaaacgcgaacgcgcg	2410
Dp	2401	AGACGACTACGACGAATCGATGCTTCACAAATTAATCATGATFANGAACGACCAACGCCCC	2460
Oy	2411	gacttcgcgaacgcgacgctgtgcacagacgcgtgtgcgtgtcgcgagctcgcgacgcgcgcg	2470
Dp	2461	GGTGTGGGTGAGAGGGACCTGTGAGCAGAATTCGAGATCCGCGAGATGTGGGGGGGG	2520
Oy	2471	caagctacatgcgtcgtgtgtcgcacgcgaatcgaacgcgcgtctcaacacaaagtgc	2530
Dp	2521	GCTGTGGTGGGTGCTCCTGTGGCGCGGACGACATTCGACGAGCCGCTCTCAACCAAGGTGCG	2580
Oy	2531	ctaacatatacaactagaaacaaacactcccgaggtattaggaagatgtgacttcaaac	2590
Dp	2581	ATAACACCAATTTCCCTCTGTGAAGACACAGATTAAGGACCTGATAAATGATGCACTTCTCAC	2640
Oy	2591	tgctgcacaattgaggtgtagatgcgaacgagcgatgcgtgcgaacgcgaacccgtgccttca	2650
Dp	2641	GGGTCAAAATTTCCGTGAACACAAAGCGGGCCCATCGACGGGATACTCTCTCAACGCTTCCA	2700
Oy	2651	cttgtactaacacgctcatcgtctagcgacgagatgcttcctcgaaacacgcgacgcgaagtgtcc	2710

Db	2701	CCTTACTATACAGTGGTGGCTGAGTACCGAGATGCTCGACAGAAATCTCTGCAAGATTTGCC	2760
QY	2711	tcacagatccaaactacttcgaagttccagcgatalcgaatagaaatcatcagacaaa	2770
Db	2761	CCCTGACCCCGACTATTATGGGAAACCGAAGAAATATCAAAATCCACATCCAGCACGAA	2820
QY	2771	caaaaagtgcctcgaagccgtctcaactcgaagaatlcacacagaaagtgatgcgtcggagaa	2830
Db	2821	CACAAAGTCCCGAGGCGGAAACGACTAAGTTTCGATACCCTGCTGTATATTTACGAA	2880
QY	2831	tycaagagcgcgagcgagggtgtgccagctgtactcccgccagccgtgcgaagacgaat	2890
Db	2881	CGCAACCCACTTACACAGAGTGGTCACTCGATAGCCAGATGTATTGACAGAGAGAAAT	2940
QY	2891	gtaccacaagttacgatacacagaatgaactttgcgttgaaccccgccagtgcggattct	2950
Db	2941	ATACCAACAGGTGAGTACGCTCATCAATTAATGACAGTAACCTCGACATGATCAACTCTT	3000
QY	2951	cgaagtgcacctgagacactgtgcgcttcgaagtgtcaatcccg-----ggagga	3001
Db	3001	CTCCGTAAACCGAAGAGCCGCGCTGGGTGATACGGACATYAGACCCAGGGATGGGGGA	3060
QY	3002	aaaattgagccgcgaatggagatagccctaacataactactttgtlaaatttcatgcaa	3061
Db	3061	GGTGTGAGCCGTGATGTGTATATACCAACGACCGTATCTTCTTCACACTCATGACAA	3120
QY	3062	ctctcttcttgatgtgtagcgttagagaaaccagagcggaattgaaatatgttcgtctc	3121
Db	3121	CTTCATGGGGGAAGAGAGAGTAAAGAAATCAGACGACAGAAAGTTCTGTATCTT	3180
QY	3122	attgtagttagaagacaagccctccatgagatgcacattgcctgtagaacttcggtttagt	3181
Db	3181	GTTGGATGTGAATACAAATCTCTCGTAAATGGCACCGCCAGCAACTCTCTTGACATAT	3240
QY	3182	ttccgaagagcagttgctgtgtccgttgcgttgaaccaccgaatctacgcagcgagcaagga	3241
Db	3241	ATCTGAGAACTTAAAGACGGGCGTCCGTTCTTGAAACCACTATCTTCCGCCCGACCGGA	3300
QY	3242	tgaaccagagacgagacaactcgcgtgtccgtttacggaatcccttgcgaactcagatcacga	3301
Db	3301	CGAGCGGACACACAAACAATCCAGGGCGGCTACGAGATCTGAACCTC--AGCACGA	3357
QY	3302	ccgagacatcgaggtgcgcgagatctcttcaacatgatctcgattgaaataaacaactcggga	3361
Db	3358	GCGGACACTTCAAATGCCGAGCGCTGTTTGATGATACAGATCCGCAAGTCCAGGGAGA	3417
QY	3362	acttgaagccgtatgtaacttgaaggaggtatctgggcacattacgaaatatcatcgaagc	3421
Db	3418	GCTGAGAGCCGCTATGCACTCAAGGATATTGTGGGAGAGTACGCTATACATATACGGGC	3477
QY	3422	cttcgagacagcgtacccgcagcagagatgtlccaaagatlaaacacctgtgtacccgcc	3481
Db	3478	ATTGCAACCAAGCATTCGCCGAATATGTCATGACAGAGACATATGAGCTATCATCCATC	3537
QY	3482	ctcaeaactccacaacctgtgttgcgttgcgtccgcaaacccgactccgtatctcgcttc	3541
Db	3538	GTTCAACTACTACCCGCGCTAGTTCGCTTCCCGAACCAAGATGCCGTGATACGACTTGC	3597
QY	3542	taggagagcgcaacaggaagcggtcttcgtgcgaagcgtccgcaacgaattcttcgagcc	3601
Db	3598	GAGGGAAGACTCTTAATCAATGAGTTCTAGGCAAGTGAACGAGAGATTCTTGGAGCG	3655
QY	3602	gattcaagcccaagcagcagcgcgtccacgcggcgcaagtcgaacttccacgtccgaag	3661
Db	3658	GATATCGCGCACTATATCCGAGCGAGCTCCACGGGGCGTGTACTTCCAGTGGTAAAG	3717
QY	3662	aaatgaagagcgcttcaagtactttgatalaacttgaagtggagcaggaagaaatagcgg	3721
Db	3718	CGATGAGGAATCAACAAGGTACTTTCAGAGTAAACGATGGC-----GAAACCTCTGG	3771
QY	3722	gcagctatatlaagcagactttcccgagagcaatcgaacaaatctagatcaagatcg	3781

Db 3772 CTGCTGAGGTACTGCAAGCCGTTCCAGAGAGATCAGGAGTCCGATTAACATTCG 3831
 QY 3782 ggcacaggaagcgagcgagcgccgcttggaccgagcttccglttccgltgct 3841
 Db 3832 CGGTACAGACGAGGAGACCCAGGACCGCTGCTCAGGACATGACGTTGAGTGT 3891
 QY 3842 ctctgtaaccacagcgagcgaccagtgltcaagcgaaatgacgacttgcctctt 3901
 Db 3892 TTTTGTGCCACGACGAGGAGAACTAGATTGCGGTCTCTCGAACAATGCTCTTTCAT 3951
 QY 3902 cgaaggtgaagaagcgctccgtgtagagtttggactgcccgaagcgaaacctaanaa 3961
 Db 3952 AGAAAAAGAGTCCCGCATGGAAGTCTCACCAACTCTCTGACCAACATCTCAAGAA 4011
 QY 3962 cccactctggaagatgactgagcgcaagatctactcaagttattgacggcaacaga 4021
 Db 4012 CCACTCTGTGAAGACGACTGTACAGCACTTACTATCGATTATGATGGCAACGCGA 4071
 QY 4022 gggctcttctgtaactgagcaagcaagcagtcacatccctctgcaagagttgacg 4081
 Db 4072 AGGTCAATTTGGGCTGTGATCTGTTCGCAACAGCTTGTCTTAAGAAAGAGCTGATAG 4131
 QY 4082 cgaagtgagcaagcgttcaacgctgacacatcgcgagcaactcgccgacgcaatg 4141
 Db 4132 GGAACAAGTGTCCCTCCACACTCTGCAAGTGGGCGCTAGTAATCGCCGATG--GTGG 4188
 QY 4142 gatccctctgcaagacttccatctctglttcaaggtcgaatgtaagaagaagcaaccg 4201
 Db 4189 CATTTCACCTTCCTGCTTCCATCTTACTGTACTGTACCGTAGGAGGACACACCTCC 4248
 QY 4202 cccaattctgagcagacatttacaacagcgagcatttgcagcttggagcagcttggcg 4261
 Db 4249 TCCAGTGTGTGAGGAGATTTGACACCCGAGGATATTCACAGCGGATCCATCCGCGAG 4308
 QY 4262 ggaattgctactgltcaagcgagcgacacagaagaagacagaccatcagctacacata 4321
 Db 4309 AGAGCTCAGTACATTCATGACGACCCAGTCTGAGGCTCGGCATTTACTATCTATAG 4368
 QY 4322 ccgtgagacatgacgtgtagcagcagcagcagcagcagcagcagcagcagcagcag 4381
 Db 4369 CTACGATCAATGATGATGACCCAGCTGTGAGGCAATGACATCGCTTCTTACT 4428
 QY 4382 gcatgagacacagcgcttcttgcgttaatatgacacagcagcagcagcagcagcagc 4441
 Db 4429 GAACGCTCAACCGGAGTGTGACGCTTATATACAGCCAGGCGCACAGATGATGACT 4488
 QY 4442 gtccaggttcagactcagcagcagcagcagcagcagcagcagcagcagcagcagc 4501
 Db 4489 GTTCAAAATTCGAAGTCAACAGCTACACGCGCGCGCTCAGAGCCGACGACGCTCAC 4548
 QY 4502 agtctaccatcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4561
 Db 4549 CGGTACGTTGATCTCTCCAGAACCGCTTACTTCTGTTCTGTCACACGCTGCACAA 4608
 QY 4562 cgttgcagcaagaagaatttcatagcgccagcttgcagcagcagcagcagcagcag 4621
 Db 4609 GGTCAAGACACAGAGACTTATTCGCGACACTTACGCGCTGGGTTTAACTGACTG 4668
 QY 4622 caacatcgaacaggtgtgcttccagcaga--caagcagcgctgtgagcagaagcagc 4678
 Db 4669 CAACATCGACCAAGTGTGCTCCGCTAACGACCCGCTACCGGCTGTGAGCAGAG 4728
 QY 4679 cagcaggtgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4738
 Db 4729 CACGCGAGTACGCGGCGCTTATACGAGCAACGTAACCGCTACTGCTGATGAGATAGA 4788
 QY 4739 ggcgcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4798
 Db 4789 ACAGATCCGATGAGCTGATCTCTGAGCTCGATACAAACACCTGCGCGCGAGTC 4848
 QY 4799 cctgtgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4858
 Db 4849 GCTGGTGTGACGAGACTTGTGACCAACTCCAGCCCGGAGCTGGGCGCTGACCTGAGCC- 4907

QY 4859 gatcgcttcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4918
 Db 4908 --TCAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4965
 QY 4919 gcttcgcagcttctgtgtagcaagaagcagcagcagcagcagcagcagcagcagcagc 4978
 Db 4966 ACTGCTTACCTTATCATGACGACTAGAGCGCTTAACCGAGCGGTTGGAAACCCCTGTGAT 5025
 QY 4979 gacgaagtaagcgtgtgtagcagcagcagcagcagcagcagcagcagcagcagcagc 5035
 Db 5026 GACGAAGTACGCGCTCACTGACTGCTGATTAAGAACCCGCGGATCGCCGCCGAC 5085
 QY 5036 caacgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5095
 Db 5086 CACCAACACACTGTGAGAGGCTTCAACCTTATCTTCAATGAAGCAATTAAGACGCGAGA 5145
 QY 5096 ctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5155
 Db 5146 TTTAGATGCAATTAAGGAGGCTTCAACGACCTGATGATGATGATGATGATGATGAT 5205
 QY 5156 gcaattccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5194
 Db 5206 GCACCTTGGCAACGCTCTTCATGATCTCTGAGTGAACGA 5244
 RESULT 3
 AB026260 5469 bp mRNA INV 24-APR-1999
 LOCUS Bombyx mori btr175 mRNA for cadherin-like membrane protein, complete cds.
 ACCESSION AB026260
 VERSION AB026260.1 GI:4666281
 KEYWORDS cadherin-like membrane protein; Btr175; btr175.
 SOURCE Bombyx mori larva midgut epithelium cDNA to mRNA.
 ORGANISM Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditypsia; Bombycoidea; Bombycidae; Bombyx.
 1 (sites)
 Nagamatsu, Y., Toda, S., Koike, T., Miyoshi, Y., Shigematsu, S. and Kogure, M.
 Cloning, sequencing, and expression of the Bombyx mori receptor for Bacillus thuringiensis insecticidal CryIA(a) toxin
 Biosci. Biotechnol. Biochem. 62 (4), 727-734 (1998)
 98276887
 2 (bases 1 to 5469)
 Nagamatsu, Y.
 Direct Submission
 Submitted (16-APR-1999) to the DDBJ/EMBL/GenBank databases.
 Yasunori Nagamatsu, Hiroshima University, Faculty of Applied Biological Science; 1-4-4 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8528, Japan (E-mail: nagamats@ipc.hiroshima-u.ac.jp, Tel:+81-824-24-7926, Fax:+81-824-22-7067)
 FEATURES
 source
 1..5469
 /organism="Bombyx mori"
 /db_xref="taxon:7091"
 /cell_type="epithelium"
 /dev_stage="larva"
 /tissue_type="midgut"
 109..5256
 /gene="btr175"
 109..171
 /gene="btr175"
 109..5256
 /gene="btr175"
 109..5256
 /gene="btr175"
 /standard_name="cadherin-like membrane protein"
 /function="receptor for B.thuringiensis insecticidal CryIA toxin"
 /codon_start=1
 /product="Btr175"
 /protein_id="BAA77212.1"

[illegible]

OY	417	gagcgtacatcgtcgaaagcttaacatcaagggtccacaagc-----ggtcgtctgcg	470
Db	373	GAAGTACCACATTCGCTAAACATCAATTACCCGGGCTCCAAATGTACATCAATTAGGCACCA	432
OY	471	atttgttaagtcgaagccaccagagccagcttgaggccctgagtttcgaacaagatgaagcagac	530
Db	433	TTTCTGCTGGTTCTTTTCATCTCACTCGTCGCCGTTATCCGACGATTAACAGATAGTAAT	492
OY	531	ggccaatgagagccttgltattacgcgaagaacaagactacgagacagcaacaatcagacgc	590
Db	493	GGACAATGTGCACTTATTTATTTACTCGAGACAAAGATTACGAGACTCCGGAAATGCAGACG	552
OY	591	tatgttttccaataccaagtgaggggtgaaatcacaagccgctaaactgltgctgttgagata	650
Db	553	TATGTTTTCACATCAAGGATAGATGGGAGACATTGGTCCGGGTGTTTCACTGCTGATT	612
OY	651	gtcaacatcgaagcaaatccgcccatactctgcaagltgttaagcgcctgctaaattccagaa	710
Db	613	GTCAATATTGTAGCGACACACGCCATTAATTTCAAGCACTTGAACCGGTTCAATGTAGCAA	672
OY	711	catgtcgagagctagactcgcagccagctcggtgtaccaaagtctgaagccgcgaggtgaatc	770
Db	673	CTAAGGGAAAGCTCTCTGACTGTAGGTGGATTAAGTGTAAACGATCCGATGGCCGATC	732
OY	771	agcaacccgtcttcaagagcttccggtgtcgaacagcagcagggtcgcaagatgaacatcttc	830
Db	733	AGCACCCAGTTCAATGAGATTAGGATCGACAGCGATCG---AGGAGACGACAAAGATCTTC	789
OY	831	tacatgltgtgaagatacgaaccccaagcagctggttcaatatgaagatgaactgttggagtc	890
Db	790	TACATTCCAAAGGGCGGAACATTTCCGGGGGAATGGATCCGATGTACATGTACCTTGGCATC	849
OY	891	aatcgcccttgaactctcgagaacaactcagcttcataattttagcgtcaacgcttcac	950
Db	850	AATGACCCGTTGAATTTTCGAGACTCAATCCACTACATATTTCTCTGTGACCCGCTTTGGAC	909
OY	951	tgcgtacccgaacaacaacacagcgtacacatgatagggtgcaagtggagaacgttagactcgcg	1010
Db	910	TCTCTCCCAAACTCAGCTACACTGTAAACCTTATGTGCAAGTGAAGATGTGGAAATCGGG	969
OY	1011	ccccctgcgttgaggagatctctctcgaigtgcaagatttgaacgaagaagctaaatcgaagc	1070
Db	970	CCGGCCAGATGGGTAGAGATCTTTTGCTGTACAACAATTGCATGAGAAACCTCTCGAGAC	1028
OY	1071	tctctccctccgcgagatagacgcgggaacgcggataatagaggccataccataccctc	1130
Db	1030	TTCCCTGTAAAGAGCTATTCGACGGAGATACAGGATCAACAACTATCCACTACAGACTG	1089
OY	1131	atcagagatgaagctgcagactcttt-----ccctggaggttgattgaagaagga	1181
Db	1090	GAACAAGCAAGAAAGATACGTTTTCACATTCGACCATCGAGGGGGGTGTAGCGCG	1148
OY	1182	gcatctcgcagcgtactgagatgcagccgcgaagcttgaagagagctttcaacctc	1241
Db	1150	GCCTATTTATACGTAGATCCATATGAGACAGGGAACAATTTCMAAGAGAACTGTTCAATTA	1209
OY	1242	accatcgttgttacaataactactagcgtagctttgcaacagagggccacaatttccatc	1301
Db	1210	TCCATTAATCCGGTATTAATATCGACAAACGAATCTTCAGCTACGCGCACCCACTGTCATA	1269
OY	1302	atgtctcaagcgcgtcaatgtatcgaagcccgagccgctgcataaagaataacagatgtgat	1361
Db	1270	ATAGTCAACGTTATTAACGACCAAGCAAGCAACCTTTGTTCAAMAGATGTGTTTGAC	1329
OY	1362	atcataggaggaanaatccaatgaactctaaacttcaaatgaagaatlttgattccatgatcga	1421
Db	1330	AATAATGGAAGAGACAGCTTTGAGAGTTGAATCTTGATCAGAGAAATTCGATTTCATGATAGA	1389
OY	1422	gatttgggtgaanaagcgtcaatatcaacagtvgaaacttgaggaagctgtctccgcaggggcg	1481
Db	1390	GATCTGGGTCCAAACGCAACATATACCCGTGAGGCTTGAAGAGGATTAATCCACCAACCGG	1449

[illegible]

Db	2527	GAGAACTCATTTGAATGATCGACTTCTACACGTGGCAATACGGTGTAAAACAGCGCGCA	2586
Qy	2622	atcgatgaaagcgaaccccttgccgtcttactgttactaacgcgtatcgtcgaagag	2681
Db	2587	ATAGATGCTGATGTCCACGGAGATPACCTTATCTACCTGTGTGGCACTGTATCG	2646
Qy	2682	tgctccctcgaaaacccgaagaaatgcttcctccagatctccaacttctcgaatltcgaag	2741
Db	2647	TCGTCACGCTGAAGACCCAGACGATGTGCCCTGAGACCCACTATTGTGGAAACTCCGGC	2706
Qy	2742	gatalcgaaatagaatcatalctcgcacaacaacaagtgccctgaagccgtctcagtgaag	2801
Db	2707	CAAGTCGATACAAATCATATGATATACCAACAAATATCTTCAACTGAACACAGACCA	2766
Qy	2802	ltcaacacgaagcgtgtacgtctcgtggagaaatgccacgaagcggcggaaggtgtgtcagctg	2861
Db	2767	TTTTAAGGGGCTGTGTATTTATTAAGAAAGATGCTGTGGAGATGAGGTGTCTCAACTT	2826
Qy	2862	tactcccccgcgcgtgacagaagaagttgtaccacaggtacgatatacagatgaactt	2921
Db	2827	ATTGGTACCGATTGGATATAGATATGATATTACACACAATTCTGTACCAATCAACTAC	2886
Qy	2922	gcggtgaaccccccgaactgycggaattcttcgaagtgagacctgagacactgtgtcccttga	2981
Db	2887	GCCGTAATCTCGTCTGTGGCGACTTCTTGCTGTGATCCGGATACGGGAAGATGAC	2946
Qy	2982	gtgcataccccggggagacgaaaaattgagccgcgatgtgggatatgagctacacatactc	3041
Db	2947	GTCATCAACACACACCGACGAGGTCTTCGACCGGGAGGGGAGAACCTTCACACAGGATC	3006
Qy	3042	tttgaatttcacatcgaatactcttcttcgatgtgtgaagctgtagagaaccagaagaa	3101
Db	3007	TTCTTTAACTCTATACGATTAACCTTTCTCCAGCAAGGAGAGGGAACAGAAACCAATATAT	3066
Qy	3102	gttgaataatttgcgtctctatcttgaatgtgtaaagaaacgcctccgtgaatgacttgcct	3161
Db	3067	GCCGAAGTACGTGTCGTACTTCTTGTGATGTAAAGACAATGTCTCAAGAGTCTCAAGAACCC	3126
Qy	3162	gatgaactccggtttgatgttcttcgcgaagagcagttgtcgtgtgtcgtgtactccagaa	3221
Db	3127	GATGAACGTATGATGTGTGGTGTGAAAGTTTAACTTAAGGGTATCCGTCCTCAGCCACAC	3186
Qy	3222	atctagacacccgagcagagttgaacgaagaaacgaactcgcgtgtgtatcggaact	3281
Db	3187	ATATACGCACTGTATPACAGACGAACCTCGACGCGAACACTCTGGATGGGATGCCCTAC	3246
Qy	3282	ctggaacctcaagatcacagcagcagacatcgaagtgtcggatctcttcacatgtatctcg	3341
Db	3247	ATCAGCCTCAGATCGCCCAACGGGAATGTGAATGTCCAGACCTTTTCAACATATATACG	3306
Qy	3342	attgaaaaacaaactvgggaacttgaaccgctatagacttgaaggggtatvtgggaact	3401
Db	3307	ATPACAAAACGTACACAGAGAACTGTGAACCGCATGTGACTTACAGAGCTTATTTGGGAAC	3366
Qy	3402	tacgaataattcatcttgaggccttcgcacaaaggttaccgcgcagcagaagttccaacgaagc	3461
Db	3367	TATGCTATTCTATTAAAGCTTACGACCAACGGCATCTCCCAACAATGTCCAAATGAACG	3426
Qy	3462	tacacacctgtatcaccgcaccttacaacttccacacacctgtgtcgtlttccgcaaaccc	3521
Db	3427	TACGAGCTCGTGAATTCGTCTTACAACTTCCAGAGCTCGTGTTTGTGTGTTTCTTAAACAC	3486
Qy	3522	gactccgtcatctgggtcttcttagggagcgcgcacagaagcgcgcgtctcgtggaagct	3581
Db	3487	GGAGACTACTCTTCACACGGCCAGGGAGAGGGCTGTGTCAACGGTTTCTGTGGCAACGTG	3546
Qy	3582	gccaaagatctccgtggagccgatctcagccacccagaggaagcgtctcagcggggagc	3641
Db	3547	GATGTGAATTTCTTAAACCCATCTGTGCTTACCGATGAAGAGCGGCTTCATGTCTGGACA	3606
Qy	3642	gtacagtttcaacgtccagagggaatgtgagagcgcgttcatgtacttgcataatactgaagt	3701

```

Db 3607 GTTGCCCTTGAAGTGGTGGTATGATACGAGCCGTCGATTTATTCACAT-----TGTG 3660
QY 3702 ggaagagaagaataagcggcagcttatatacgcgaacttctccagaggaacaatcaga 3761
Db 3661 AACGACGGGGAACAAAGTGGCACTGTGATGCTGAACAACCTCTCCCGACGACATTAAGA 3720
QY 3762 caattcaagatcacatccggccacagcggcgacggcgccggccggcgttggacc 3821
Db 3721 GAATTTGAGGTGACGATACGCTCTACTGACGGCGGACAGAACCAAGACCGCTCCACG 3780
QY 3822 gacgcacagcttctggtgtctctctgaccacacagggcgagccagtgctcagcgaat 3881
Db 3781 GACTCACCTTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3840
QY 3882 ggaactactgtcgcctctctcgaaggtgaagaagcctcgcgtgagaagttttagctgcg 3941
Db 3841 ACTCACACTGTTGCTTCTTCTTGAAGAAAGAGGGGCTTTTGGACGTCATGAGCTGGCG 3900
QY 3942 caagcagaagaccttaaaaaaacacactctcgaagatgacttcgaagatactactaagg 4001
Db 3901 CGTGTGAGGATCGGAAAACCATCTTTGCAATGATGATGATGATGATGATGATGATGATG 3960
QY 4002 ttatttgaggaacaacagaggtcttctctgacttgaccagctcgaagcagctcacc 4061
Db 3961 ATTATTGATGAAATATATGACGGCCACTTTGACATCGAAGAACTGACAACTGCTTATTC 4020
QY 4062 ctctgcgaagatctgacccggcaggtgacacgctcttaccgctgacacgctgcggagc 4121
Db 4021 CTGGTAAAGGATTTGGATGGAGTGCACGCAACGTTATACATGACGATGCGTGGCAGC 4080
QY 4122 aactgcgcgcacgcactctggaatccctctcgaacttccatccctctgctcagctcaat 4181
Db 4081 AACTCAC-----CGACAGCGGAATCGCTTAACTTACATACATACATACATACATACAT 4134
QY 4182 gtaagaagaagcgaacccggcccaatttctcgaagcagacttctacacagcggtcattcg 4241
Db 4133 GTGAGAGAGGAGATCCCGACGCTGATCTGTCAGAGATTTGTCAGGCTGGCATTTTC 4194
QY 4242 acgttggaagcagcttgccgggaattgtctactctgacggcgagccacacagaagcagc 4301
Db 4195 ACTTGGATTCATACACAGAGACGATGTTGATTTTACAGGCGACACATTCAGAAAATGCA 4254
QY 4302 accatcacgtiacacatatagaccgtctgacgacgtcgtgacagcagccttagaagcgctg 4361
Db 4255 CCGATATATCTACACGATAGACTGTCGACGATGATGATGATGATGATGATGATGATGATG 4314
QY 4362 cgcgactgcgctcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4421
Db 4315 AGGAGACAGCCTTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 4374
QY 4422 accgcttcacatgacggcagcttctgagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4481
Db 4375 ACTGCGTCATACAGGAAATGTTGCAATTCAGATGCTGCTACGATCGATCGATCGATCG 4434
QY 4482 atcgcacacacccgctgtaagcttaccatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4541
Db 4435 ACTGATAGAGCAAAATGTCAAGATCTACTGATCTCCACAAAGAACCGGCTCTTCTTCTTG 4494
QY 4542 ttgcataacaaacttgagacgtctgacagaaagaattcatatagcggcagctcagc 4601
Db 4495 TTTGTTAACTTTAGAACAAAGTGCAGAACACTGACTTATACACAAAGTTTCC 4554
QY 4602 accggttttaacatgacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4661
Db 4555 GCGGCTTGAATGACGTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 4614
QY 4662 gtgagcagaacacacacagcaggtgctgcgcgacttcatccggagacagctgcgcgtg 4721
Db 4615 GTTCATCATTAAGGATCACCAGAGTGAAGGAGGACACTTATCCGCAATATATGTCGCCGTA 4674
QY 4722 cagcagaagaagctcgaagcgttccgacagcagaacggtgctgctgcgcacatccagctg 4781
Db 4675 CCGCCGACGAGATGAGACTCTCCGAGGGGACATGTTGCTGCTGACTGCTCCTCCTCCTCCT 4734

```

```

QY 4782 atcgtgaccacacagcctgtgtcgtcgaagacgtgtgacggtgtacactccagccta 4841
Db 4735 ACGCTGGCAGCGCGGCTGTAGTGTGCGGACGACTTTCACGAGACACAGTCCGGCGCC 4794
QY 4842 ggcgaagagatcaatgacatgcgcgtctcgaactagcgcgcgtcgcgtgtgtcgtgagc 4901
Db 4795 GACGCGGGCTC---GGCCCGGCTGTGTACGCGCTGCGCTCTCTCACCCGCTGCTGT 4851
QY 4902 ttccctgcgtcgttactgtcttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4961
Db 4852 GCGCTGTGTCTCTACTGCTAGTCACTTCAATCATCATCATCATCATCATCATCATCATCATCAT 4911
QY 4962 ctgcaagcactctcactgacgaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5021
Db 4912 TTGGAAGCCCTGACGGTAAAGATGACGATGCGTGGATTCAGAGTTTGAACCGCGTGGC 4971
QY 5022 ct---ggcgcgggacacacacagcagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5078
Db 4972 ATCGCGCGCGCGGACCAATAAACATGCGTGAAGGCTCCAAATCCGATCTGGAACGAG 5031
QY 5079 ggcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5138
Db 5032 ACCATTAAAGCCCCGACCTTCGATTTATGAGGATGATGATGATGATGATGATGATGATGATG 5091
QY 5139 ggcacgcgaagacatgcgcgaatccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5194
Db 5092 GGGATTGAAGATGCTGCCACACTTCGCGCAAAATATATTTTCTGTCGTGCGTGA 5147

RESULT 4
AB041508 5172 bp mRNA INV 11-JUL-2000
LOCUS AB041508
DEFINITION Bombyx mori btr175a mRNA for cadherin-like membrane protein,
partial cds.
ACCESSION AB041508
VERSION AB041508.1 GI:9049394
KEYWORDS Btr175a.
SOURCE Bombyx mori larva midgut cDNA to mRNA.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
1 (sites)
Ikawa,S., Tsuda,Y., Fukuda,T., Sugimoto,K. and Himeno,M.
cDNA cloning of the CryIaA receptor variants from Bombyx mori and
their expression in mammalian cells
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 5172)
AUTHORS Ikawa,S., Fukuda,T., Himeno,M. and Sugimoto,K.
TITLE Direct Submission.
JOURNAL Submitted (07-APR-2000) to the DDBJ/EMBL/GenBank databases. Takashi
Fukuda, Osaka Prefecture University, Department of Applied
Biochemistry, 1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan
(E-mail:tfukuda@biochem.osakaifu-u.ac.jp, Tel:+81-722-54-9464,
Fax:+81-722-54-9463)
FEATURES
Source
1..5172
/organism="Bombyx mori"
/db_xref="taxon:7091"
/dev_stage="larva"
/tissue_type="midgut"
1..29
/feature="PCR primer site"
25..5172
/feature="btr175a"
25..5172
/feature="btr175a"
/feature="receptor for B. thuringiensis CryIaA toxin"
/codon_start=1
/product="cadherin-like membrane protein"
/protein_id="BAA99404.1"
/db_xref="gi:9049395"

```

```

/translation="MGVDVRIATLTLIIYAETVLAQERGFVNAIPRPDPPLDELDF
EGOTSORBILPADREDVDCMGYAMPTQTOITVMEPEEGEVPIAKINRPNV
PYIPERATSGSFNLLVPIYIRIPDSNGMHLITTOROYEPTGMQYFNIRIDETL
VAGVLLIVNIDNAP110ALEPCVDLGEARLVEVYVTDAGVSTOFMERID
SDRGDKIFYI0GANIPEGMITMTVAINSEPLNEETPLHIFSVTALDSLPNTVT
LMVOYENHEHPRWELEFAVOEDEKTAQSEFPVADIGDINKPIHRLTEAEDVT
FEHITITIGRSGALIVYDPIRDLTLOREVOPLSTIAKYDNESGATANVYIYNDI
NDOREPELFEKRYNIMETALTUNFDRPFGHOGNOAYIRLESDEPDAKAA
FYIAPVEYQROQPIMTAKNMKLDYEPERQIRLRIATDMDEEHVAVYIYNL
IIMNDEEPIEHSVONVSFEKEGKFVAVNRADRIDRVEHTLMLNNNLSTID
KDQDIIHYTODDFEDYHROSELFOVNRADDLGEPFHATSQLIHLIDINNPTTLR
LPRSPNVEENVPREGYIITSEIRATDPTALELEIDMTSYATKQGRANPLEFN
CVEIETIYPA1INRGSAIGRLVKKIRENVIIDEEFMLTIVRDLNIVYDGDYD
ESTPITIIDMNDPMIVPGTLOSILVRMSAGVAVYITVDIDIDINGPLXNOYRT
MKANEGPENLIMIDFTGQITVTSAGIDYDVRVRLYITVATDSTAEDEDDCC
DDPIYETPGQVYI0IIDTNKIP0PETDOKAVVYIYEDVSDVEVVKVGSJDRD
DIHYIRQIYNAVNPRLDFEAVDPDGRVYVYITDEVDLROGDEQOHRIFENLID
NEFOGDNRNQNDAEVLVLLDVNDNAPELPEDELSMSVSESLTKGTRLOPHIYAP
DRDEPDJNSRVGAIISLTIANREIEVPELTI0IONVGETLADLWGWTYTA
IHIRAYDHGIPOKMSNETYELVIRPNHAFVFPKRGATRLRLAREAVNGILATY
DGEELNRIVATDEGLHAGOVAFEVGTEAVDFPHIYDNESGTLMLKOLPEDIR
EPEVTRATDGTGERPLSTOCITFSVFPVIOGEP1EPTSTHVAE1EKEG1ERHE
LPRAEDRNLHCSDDCHN1YRIIDGNNDGHRGDETTNVLVKELDROVSERYTL
IAASNPFGIALSTITITIVNREADQPYFVRDLVAGISTDSIRRELLI0QATP
SENAP1IYIDMSTMTDPTLASVRETAFLINPHGVTLTNIOP7ASMHGMEQVVA
TDPAGVSDRABVKKIYL1STRNRVPELNTLEOYNDPFIAPESAGEEMTCMDVY
VPADASVYINNGITEVRNGH1RDNVPRPADETITLRKDMYLA1R0STLRLVLR
DLFTDTSAPDPAGSAALVYLAVALSALIAALLL1V1L1R1K1RL1EAL1V1K1Y
G5V0SLNRVGI1AAGPKNAHVEGSDSNP1MNT1KAPDSGSDSN0SD1G1ED1LPH
EGENNYPRDVEKTDKPEDIVATHNNNFEGKSTPSPERANOFOKL"
4725..4803
/misc_feature
/gene="btl175a"
/note="putative
transmembrane region"
primer_bind
5155..5172
/note="PCR primer site"
BASE COUNT 1483 a 1292 c 1225 g 1172 t
ORIGIN

```

```

Query Match 35.9%; Score 1975.2; DB 3; Length 5172;
Best Local Similarity 63.2%; Pred. No. 0;
Matches 3207; Conservative 0; Mismatches 1823; Indels 48; Gaps 9;

```

```

144 aactaagatcatgtaaatgggttgaagagttctcccaagcagtgctactgtctct 203
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
7 AACCAAGGAACATGCAAGATGGAGTTCGAAATTCAGTACACTACTTCTAATC 66
204 ttagccttcgcgaactagccaacacgaagttcgtacattatgcaatccaagcag 263
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 TATGCAAGAACTGTTTGGCTCAAGAAAGATCGGCTTCATGTTGCCATACCCAGACCA 126
264 gagactccgaactgcgcctatttgatcaagagaaatcatggaagtgaacagcctta 323
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
127 CCAATGCGGATCTGCTGACTAGACTTTGAAAGTCAAACTGGAGTCAAAACCACTA 186
324 ataccgcgcgaacccgcgaggaatgatgtatgagaact-----tctaccgat 374
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 ATACGGGGGAGACCGCGAAGATGCTGCATGCGCGCTACACGCGCATACACCTACC 246
375 caaatgactcaggtcatatcatagaagaagaatcgaagaagacgtcatattgagag 434
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 TAGGCAACTCAGATTATTATTACATGGAAGAGATAGAGGAAGATACCATCGCTAAA 306
435 cttaaactcaagaggtccaacacgc-----ggtcgtcgtattatgtcagcgacgcc 488
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 CTCATTTACCGCGGTCCAATGATACATATTGAGCCAGCATTTCTGTGTTCTTTC 366
489 agagccgcagctgggcccgaatttcgacagaatgaagcagcgccaattgagccctgtc 548
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 AATCTACTGTGTCGCCGTTATCCGACGAGTACAGATAGTAATGAGAAATGGACCTTATT 426

```

```

549 attacgcaagaagaagactacgagagacgaacacatgcagagctatgtctcaatcaa 608
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
427 ATTACTCAGACAAAGATTACGAGACTCCGGAAATGACACAGTATGTTTCAACATCACG 486
609 gggaggggtgaatcacaagccgctactgtgtgcgtgagatagtcacaatcagacat 668
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 ATAGATGGGGACACATTGGTCCGGGGTGTTCACGTCGATGTGCAATATTCAGCAAC 546
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 ccgccatctcgaagtgtcagcgctgcgtaattccagaacatgsgagactagctg 728
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
547 GCACCTATTAATCAAGCACTTGAAACCGTGTCAAGTTGAGCAACTAGGGGAACCTGCTCG 606
729 accgactcgtgtaccagagttgcacacgcgagcgtgaaatcagacccgcctgcag 788
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
607 ACTGAGTGGGTATACGTGTGTAACGATGCCGATGCCGATACGACCCAGTTCAATGAG 666
789 ttcctgtcgaagcagcagggctcgaatgaagaacatctcacaatgttgaagaatc 848
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
667 TTTAGATGACAGAGGATCG---AGGAGACGACAAAGATCTTACATCCAGGGCGAAC 723
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
849 gaecccagcagctggttcaatgaatgactgtgggataatcgccttgaactc 908
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
724 ATTCGGGCGAATGGATCCGATGACAAATGACCCGTTGGCATCAATGAACCGTTGAATTC 783
909 gagaaactcagctcatatattagcgtcaacagcttctgactgcacggaacacac 968
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
784 GAGACTAATCCACTACATATATTCTGTGACCGCTTGGACTCTCTCCAAACCTCAC 843
969 acggtcacacgatgtgtcgaatgagagcgtgagtcgtgccccctgcgtgtgag 1028
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
844 ACTGTACGCTTATGCTATGACAGTACAGATGAGAACATCGGCCGACGATGGGTAGAG 903
1029 atcttcagtgcagcagcttgcacgaagaactaactcagagcttccctcgcgcgata 1088
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
904 ATCTTGTCTGTACAACAATTCGATGAGAAACTGCTGACACTTCCGTGAAGACTATTC 963
1089 gacgggagcagggaaataaataagggccatcaactaaccctcatcaggaatgaacgtgac 1148
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
964 GACGGATACAGGATCAACAACTATCCACTACACTGGAACGACGAGAAAGAT 1023
1149 gactctctt-----tccctgaggttgatgaaagcagagctatctcgcgtgact 1199
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1024 ACGTTTTCATATCCGACCATGAGGGGGGTGTGTACGGCGCATATTTATAGTACAT 1083
1200 gagatgcacgcgcgaagccttgaagaagagctttcaactcaccatcgtgtgtacaaa 1259
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1084 CCAATAGACAGGACACATTTACAAAGAGAAAGTGTTCATTAATTCATTAATGCGGTATAA 1143
1260 tctactgagcgttagcttgcaacagaagccacatlttcatcatcgtcaacgaagctcat 1319
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1144 TACGACAAACGATCTTCACGCTACGCGACCAACGTTGTCAATATATGTCATACGATTAAC 1203
1320 gatcagcagccgcgcgctgcataaagaatatacagttatgatalatcatgaggaactcca 1379
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1204 GACCAAGACACGAAACCTTTGTTCAAGAGATATGTTGAACATTAATGAGAGACAGACT 1263
1380 atgactcaaacctcaatgaagaatttgatccaatgatcagatgtggtgtaaaagcgt 1439
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1264 TTGACATGTAACCTTCGATCAGAAATTCGAAATTCATGATGATGATGCGGTCAAAACGCA 1323
1440 caatacaagactggaacttgaagacgtgttccgcgaagggcggtccgcacttcaatc 1499
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1324 CAATACACCGGTGAGGCTTAGAGACGATTAATTCACACAGACGCGGGAAGCCCTTTTACATA 1383
1500 gcgcgggagcgtcactaccagagcagaccttcatctatggtcaccataaaccacacatg 1559
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1384 GCACCCGAAAGTTGGCTACACGACAGACGCTTCATTCATGGGACAGGTAAACCAATAAATG 1443
1560 ctggaattacagaagatgtcatlttccagaatacatcaatlaagtcacaagcaggtgacatg 1619
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1444 CTCGATTAATGAAAGTTCCAGAAATTTCAACGAGTACAGTTGCGCGGTAAATAGCAAGACATG 1503
1620 aacaacgctagcacaagtgaggagcgctgtgtcgtgaacccgtgactggaacagac 1679

```

1504 GACATGAAAGAACGTCGGGTTGCGATATGTGATACATCAATCTTATCACTGATATGAT 1563
1680 gaattcccatcttcgaagagagcagctactccgcgttcgtttaagagagcgtcgagcc 1739
1564 GAAGAGCCTATTTTGAACATAGCGTACAAAACGTTCTCTTCAGAAAGAAAGGAGAG 1623
1740 gcttcccggtgagcagcgtgctgcgccttcgacagagacatcgaacgttagtgat 1799
1624 GGATTTTTCGTGGCTAATAGTAAGAGCCACAGACAGACATTGATGACAGAGTTGAACAC 1683
1800 tcatgtatggcgaacgcgtgttgaactcgtlcatagatgaatcaacgggagagatcttc 1859
1684 ACATTAATGGGAACGCGCAATTAATTAATCTCTCATAGATAGTAAGACACAGGGGACATTCAT 1743
1860 gtggagcagtgacgtatgcttgcactacacacgaagaacactatattgttcaagtgagc 1919
1744 GTACGAGGAGAGCTTTTGTGACTATCACCGACAAACGAACTTTTGTTCAGGTTCAGA 1803
1920 gctgaagatcttggcgagcagcgccacacacacagtgacacccagcgtgtgataaa 1979
1804 GCTGACGACACCCGCGGGGA---ACCATTTACAGCGGAGCTTCTCAGCTGCTGATTCAT 1860
1980 ctgagagatgtcaacaacacatcctccacacccctgaagcttgcacctcgaagcgtc 2039
1861 CTGGAAGATATTAACACACCCCTCTCTACTCTGAGATTGCCGCGGACGCTCTAATGTC 1920
2040 gaggagaacgttcccggaagatagagatataccggaaatcactgcgtcaccaccggag 2099
1921 GAGGGAATGTCCCGAGGCTACATCAATCAAGTAGATCCGACCGCATCTTATGAC 1980
2100 accgagcctcactcgtgtgttcgagatcagctcgtgagcctcgtgagccacgaagcagc 2159
1981 ACTAGGCTGAACTGAGCGCTTGAAGATGATCGAGCACTTCTACGCAACCAACAAAGC 2040
2160 agagagacaaacctactgaatacgtcggtgtatagttatcgaacagatatcccaacc 2219
2041 AGGGAAGCTAATCTATGAAATTCACAAATTTGTGAGATCGAAGCAAAATATATCTCGC 2100
2220 gagggaacccgggttccgcactcgagcgctgtgtgtgaagagatccggaagacgtc 2279
2101 ATCAACAACCGCTGGCTATGAGAGCGCTCGTCAAAAAAATCAAGGAGAACGTC 2160
2280 accatcagcttcgagaaattcgagatgtcttactcaccgctcgcgtgaggaagcctaac 2339
2161 ACAATTGACTAGGAAGATTGAGATGTTGTAATCGACGCTGAGAGTGGGAGCTTAAAC 2220
2340 actgtcagcagatgactagatcagatgagcagcgttcaacgatcaaatatagatgaac 2399
2221 ACTGTATCCGGGATGATGATGATGAATCAACCTTCACATCAATCATCATGATATGAAC 2280
2400 gacaacgcgcgacttctcgaaacgcgctgacgagagacagatcgtgtcgagagtg 2459
2281 GATTAATCTCCGATATGGGTCCCGGAAACACTGAGACAGTGCCTGCTGCTGGGAAATG 2340
2460 gcggccagcagcagcgtcactcgtcgtgtcgcacacgaacatcgaacgcgcgtctac 2519
2341 TCAGACGCGGTCGTGATTTGGGACTCTGACGGCCATGATGAGAGGACCACTTAC 2400
2520 aaccaagtgagctacacatatacaactagaacaacactcccgagagattagtgagatt 2579
2401 AATCAAGGCAATACATATGAAAGCAAAAGACAGACACAGAAATCTTATTTGATGATC 2460
2580 gactcacaacgtgtcaaatgtgaatgtgagtgcaacgagcagatcgtcagagacacc 2639
2461 GACTTCTCACTGCTGATACATACGCTGAAACAGCGGCGCATATGATGCTGATTTCCA 2520
2640 tggcgtcttactgtactacacgctcactgtagcgaagcgtgtccctcgtgaaacgc 2699
2521 CGGAAATACAACCTTACTACACTGTGGTGGCACTGATGCTGATGAGTGAACCA 2580
2700 acggaatctctccagatccaaactactcgaagttcgaagcgatatacgaatagaatc 2759

2581 GACGATTCCTGACGACCCGACTATTGGGAACCTCCCGGCCAAAGTCGTATCAAAATC 2640
2760 atcgacacaaacaaagtgctgtagccgcgtcactgagaagtlcaacagcagtgatc 2819
2641 ATAGATACCAACCAAAATATACCTCAACCTGAAACAGACCAATTAAGCGTGTGAT 2700
2820 gtctggagaaatgcacagcagcagcaggtgtgtcagctgtactccacgaacgtgac 2879
2701 ATTTCGAAGATGCTGTGTCTGAGATGAGGTTCGAAGTATATGTAAGCATTTGGAT 2760
2880 agagacaggtgttaccacacaggttaacatacagatgaactttcggtgaaccccgatg 2939
2761 AGAGATGATATTTCACACAAATTCGTAACAAATCAACTACGCCGTAATCTCTGTG 2820
2940 cgggaattcttcgaagtgagccttgacactgtgtgccttgaggtgcatcacccggagac 2999
2821 CGGACTTCTTGTGCTGTGTGATCCGATACGGGAAGGTGATCTACTACACACGAC 2880
3000 gaaaaattgacccgcatgtggatgagcttaccacatactatcttgaattcatgac 3059
2881 GAGGTCTCGACCGGAGCGGAGACGAACTCAACAGAGATCTTCAACCTCATTCAT 2940
3060 aactcttcttctgtagtgagcaggttagagaaacagcagcgaagttgaatatgttgct 3119
2941 AACTCTTCCACCAAGAGAGACGCAACAGAAACAAATGATGCCGAATCTGTGCTGA 3000
3120 ctatggatgtgaacagacagcctcctgagatgpcattgtccctgagtaactccggttgat 3179
3001 CTCTTGTGATGTACAGCAATGCTCCAGAGCTTCAGAACCCGATGMACTGATGCTG 3060
3180 gtccgaagagcagtgctgt 3239
3061 GTGTCTGAAGTTTAATAAGGATACGGCTCCAGCCACACATATATACGACCTGATAGA 3120
3240 gatgaacagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3299
3121 GACGAACCTGACACGCAACACTCTCGGTGGATGCGCATATCATGACCTGATGCC 3180
3300 gaccgagacatcgaggtgtccgagatcttcaacatgactcgtatggaacaaacactgg 3359
3181 AACAAGGAATGGAAGTCCAGAGCTTTTACCATGATATACAGATCAAAACGTCACAGA 3240
3360 gaacttgaacccgtatgtacttgaagggtatgtgggcaacttgaacaaatattcatag 3419
3241 GAACHTGAACCGCATGAGACTTGAAGGCTATTTGGGAACTTATGATATCATATTA 3300
3420 gcttcgacacacgttaccgcgaagagaggttcaacagcagagatccctgtgtatcgc 3479
3301 GCTTACGACCAAGGATCCCGCAACAGATGTCATGAACCTTACGACCTCTGTATTCGT 3360
3480 ccttcaacttccacacacccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3539
3361 CCTTCAACTTCCAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
3421 GCCAGGAGAGGCTGT 3480
3600 cagatctacgcacacgaagagagagccttccacgcggtgcaagtgactgttccagtcag 3659
3481 CGCATCTGTCTTACGATGAAGAGCGGCTACATGCTGCAAGTGGCCCTCAATGCTGT 3540
3660 gaaatagaagagcgttccagtgacttgaataactgaagtgagagcaggaataatagc 3719
3541 GGTGATACTGAAGCGTGTGATTTTCCACAT-----GTCAAGAGCGGAGAGAACGT 3594
3720 gggcagctataatagccagcttcccgagcaaatcgaacatcagatcagatc 3779
3595 GGCACCTGATGCTGAACCAACTTCCCGAGACATTAAGAAATTGAGTGAACATA 3654
3780 cgggcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3839
3655 GCTGCTACTGACGGGAGACAGAACCAAGACCGCTCTCCACGACTGCACCTTCAGTGTG 3714

[illegible]

QY	4920	ctctgcgcatgtgtctctgttagaacaaagacacgacgaacggcagctcgaagcatctccatg	4979
Db	4786	CTACTGATCTTCATCATCATCAGAGACTTAAAACTTAAACCGCTGTTTGGAACCCCTGACGGTT	4845
QY	4980	acgaagtagcgcgtcgcgtgtgcactccgcggctgtbaaccgcgcgggcgt--ggcgccggagcac	5038
Db	4846	AAGAAGTAGCGGATGGGTGGATTCAGGTTTGGAACCGCGTGGCATCGCGGCGCGGGACAC	4905
QY	5037	aacaagacgcgcgttcgagggcttccaaccccatgttgaaagagagcatccgcgcggccgac	5096
Db	4906	AATTAACATGCGCGGGAAGGCTCCAAATCCGATGTGAAAGAAACCATTTAAAGCCCCGAC	4965
QY	5097	ttcgacgcacatcagttgacgcgcgttgccgcactccgacactatgcgtcgcagagacatctcg	5156
Db	4966	TTTGATTTCTATGAGCGATGCATGAACAGACTCGATCTAATCGGCAATTGAAGATCTGCCA	5025
QY	5157	caattccgcgacgactaactcccgcccgccgacacaga	5194
Db	5026	CACTTGCGGCAAAATATATTTTCTCTCGAGAGTGTGA	5063

RESULT	5
AB041510	
LOCUS	AB041510 5172 bp mRNA INV 11-JUL-2000
DEFINITION	Bombyx mori btr175c mRNA for cadherin-like membrane protein, partial cds.
ACCESSION	AB041510
VERSION	AB041510.1 GI:9049398
KEYWORDS	BCR175c.
SOURCE	Bombyx mori larva midgut cDNA to mRNA.
ORGANISM	Bombyx mori

REFERENCE

AUTHORS Ikawa,S., Tsuda,Y., Fukuda,T., Sugimoto,K. and Himeno,M.
TITLE cDNA cloning of the CytIIa receptor variants from *Bombix mori* and their expression in mammalian cells
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 5172)
AUTHORS Ikawa,S., Fukuda,T., Himeno,M. and Sugimoto,K.
TITLE Direct Submission
JOURNAL Submitted (07-APR-2000) to the DDBJ/EMBL/GenBank databases. Takashi

FEATURES

Source

primer_bind

denes

and

```

1. 5172
/organism="Bombyx mori"
/db_xref="taxon:7091"
/dev_stage="larva"
/tissue_type="midgut"
29
/notes="PCR primer site"
25. 5172
/gene="btlr175c"
25. >5172
/gene="btlr175c"
/function="receptor for B. thuringiensis CryIaA toxin"
/codon_start=1
/product="cadherin-like membrane protein"
/protein_id="BA99406.1"
/db_xref="GI:9049399"
/translation="MVDYDRIALTLILLIYIAETVLAQRCGFMAIRPPRPDLPEDLP
EQTQTSQRLPIAPADEDEDVCMQSYHAMTPPYGTQIIYMEETIEGVAIARLANKRGPVNV
PIEPAEFLSGSNLALPVYRIAPDSNGEWHLITTORODEYTPMOOYVFNIRIDGELT
VAGVSLIYINIDNAPIIDALEECQYDGEALTEGCVYVTDADGIIISOPMOFRID
SQRGDGIKIEYOGANIPGSMIRMTKQVNEPIENFENPLHITSVALADISLPMNHVY
LMVQYENHRRPRRWVEIFAYVOQFDEKTRQSPFYRAIDGTPGINKRIHYLLETRAEEDT
FPHIRYIBSGRSGLIIVPIDDTLQREYFQSLIAYKIDNESSAFANAVVLIYNDI
NDQREPLEPKEKRLNIMETBALINFDQEGFDRDLGQNAQVYTRLESQYPADAAEA
IETAEVGEQGTQETINGTANHKMLDIEVEFQRIIRLIVATIDMNDNEHVSAVYINL

```


Db	1684	ACATTAAAGGGAAACCGCATTAATCTTCGATATGACAAGACACAGGGACATTTCAT	1733
Oy	1860	gtgagcacygacgtatgccttcagactacccaacggaacagacactctatctgttcaggtgcg	1919
Db	1744	GTCACGCGAGCGACGACTCTTTTGACATATCACGCCGACAAGGAAGAACTTTTGTTCAGGTG	1803
Oy	1920	gctcagcatactcttggcgcgacgcccacacaacagctggtaccacccagctgtgtatgaa	1979
Db	1804	GCTGACGACACCCCTGGGGGGA---ACCAATTCACACGGCAGCACTTCTCAGCTACTGTCAT	1860
Oy	1980	ctgagagagtgcacaacacactccctccacccctcagcttgcgcccgttcgactcaagcgtc	2039
Db	1861	CTGGACGATATTAACACACACCCCTCTACTCTGAGATGTGCGGGGCGACGTCTAATGTC	1920
Oy	2040	gaggagaaacgtctccggaagatagcagatactcccggaatacatctgctccgaaccgagac	2059
Db	1921	GAAGGAGATGTGCCCGAGGGCTACATCATACAGAAGAGATCGAGCCACCGATCTCTAC	1960
Db	2100	accagcgctactcctgtgttcgcagatcgcagctgggactccaccttggcgcaacgaagagc	2159
Db	1981	ACTACGGCTGAACTGCGCCTTCGAAATCGATGGACCACTTCTACGCAACCAAGAC	2040
Oy	2160	agagagaccacccctactgatatcgtcgggtgatatgtatcgaaagatataccaccac	2219
Db	2041	AGGGAGCTAATCTATTTGAATTTCACAATTTGTGTGAAGATGAAACAATATATCTTCG	2100
Oy	2220	gaggagcaacggggtctccgcacatcgcgccctgtgtgttcaagagatccggagcaagctc	2279
Db	2101	ATCAACAACCGTGGCTGGGCTATAGAGCGGCTGTGTCTAATAAANAATCGAGAAACCTC	2160
Oy	2280	accatcgactctgaggaattcgcagatgcttacctccaccgctcgctgagggacctcaac	2339
Db	2161	ACAATTTGCTACGAAGAGTTTCGAGATGTTGTATCTACCGTCAGAGTGGGGACCTTAAC	2220
Oy	2340	acttcactcgagatgtactcgcgtatggaggaagttacagatacaataatgcagatgac	2399
Db	2221	ACTGTATTCGGGGATGTATATGATGAATCAACCTTCACATACATCATGATATGTAAC	2280
Oy	2400	gacaaacgcgcgaatccttcggaacggaacgctgacgcaagaatgcggttcgcgagctg	2459
Db	2281	GATTAATCTCTCGAATATGGGTGCCGGGAACACBTBAGACATCCCTCGTGTTCGGGAATG	2340
Oy	2460	gcggcgcaacgcgacgctcatcgcgtctccggtctcgcgcaacgcatcgacgcgcgctcac	2519
Db	2341	TCAGACGCCGCTGTGCGATTTGGGACTGTGAGCGCCACATGATGAGAGCACTCTAC	2400
Db	2520	aaccaagtgcgtactacataataaactagaaacaacactcccgaggatagtgaaagatt	2579
Db	2401	AATCAAGTGGGATACACTATTAAGAGCAAAACGAAGGCACACAGAAATCTATTGTATGC	2460
Oy	2580	gacttcacaactgttcaaatatgtggtgtgagtgtgaacagagtgatctgactgcagcaagacc	2639
Db	2461	GACTTCTTACACTGTGTCAGATCACAGTGAATAAGGGGGCAATATGATGCTGATTTCCA	2520
Oy	2640	tgagcgtcttactatgtactacacgcgtcatcgctctagcgagaaatgtctcccttgaaaacgc	2699
Db	2521	CGGAGATATCAACCTATATCTACACTGTGTGGGCCACATGATCGGCTCTACGCTGAAGACCA	2580
Oy	2700	acgggaatgctccctcagatcttccaactactctcgagttccaaggcgatatacgaatagaattc	2759
Db	2581	GACATTTGCCCTGCAGCACCACTATTGTGGGAAACTCCCGGCAAGTGTGATCAAAATC	2640
Oy	2760	atcgacaacaacaagaagtgctcgcgctcgcctcacttgaagttaacaagaagtgctac	2819
Db	2641	ATAAATATCCACAACAATAATACCTTCAACTGAACCGACCAATTTAAAGCGCGTGTGAT	2700
Oy	2820	gtctcgggagaaatgcacagagcgcgacgagtggtctcacgctgttactcccaagacgctgac	2879
Db	2701	ATTATACGAAGTGTGTGTGCTGAGATGAGGTGTCTCAAAAGTTATTTGTAAGGATTTGGAT	2760
Oy	2880	agagagagtgcttgcacaacacggtgacgatacacagatgaaacttgcggtgacccccgactg	2939

D	2761	AGAGATGATATTATTACACACACATTTCTGTACCAATATCACTACCGCTGATCTCTGCTG	28220
O	2940	cggagattctcttcgcgagtgtygacacttgycacatgctgccttgagtgatccacggggagc	29999
D	2821	CGCGACTCTCTCGCTGTGTGATCCGGATACGGGAAGAGTGTAGCTACTACACACCAGC	28808
O	3000	gaaaaattggaacccgcgatgggatgagcccaacataactactctttgtaattatcatgat	30558
D	2881	GAGGTCTCGACCGGGACGGGGAGAACCTCAACACAGAGATCTCTTTTAACTCTATCAT	29400
O	3060	aactctcttctcgtatgtgttcacgtatagagaaacacagacgaagtgaattatcttgcgt	31199
D	2941	AACTTTTTCCAGCAGAGAGACGGCACAGAAACCAAAATGATGCCGAATACTGTGGTCTA	30000
O	3120	ctatbtgagtgtgaacgcgaacacgcctcctgagatgcattgctcgtatgaactccggtttgat	31799
D	3001	CTTCTGTGACGTGAACGACAAATGTGCTCCAGACTTCCAGAACCCGATGAACTGTGCGTGC	30600
O	3180	gtttccgaaggagcagttgctgtygtcctgtactcccgaaactcagccacggagacag	32399
D	3061	GTGCTGAAATTTAACTTAAGGGGTAGCGCTCTCCACCCACACATTTAGCGACCTGATAGA	31200
O	3240	gatgaaccagacacagagcaactcgcgtgtcctgttctagagaatccctgagactcaagatacc	32999
D	3121	GACCAACCTGTACACGGACCAACTCTCGGGTGGGTAGCCCATATCAGCTTCACGATGCC	31800
O	3300	gaccgagacactcggagtgycgcggaatcttcacatgatacttgatgaaacaaacatctgg	33599
D	3181	AACGAGGAATCGAAGTCCAGAGCTTTTCACATATACATACAAACGTCAACAGA	32400
O	3360	gaacttgaacgcgctatbgaacttgaaggaggtatbtgggcacttacaataatcaatgaag	34199
D	3241	GAACTGGAACCCGCAATGGACTTGAAGAGGCTATTGGGGAACTTAATGCTATTAATTAA	33000
O	3420	gccttcgacacagcgcctaccgcgcagcagagtgccaaagagacgtataccctgctatccgc	34799
D	3301	GCCTACGACACACGGCATCCGCAACAGATGTCCAAAGAAACGTACGACTGTGATTTGCT	33600
O	3480	ccctacaactctccacacccctgcttgcgtgttccgcgaacccgcgaactcgcattccgctt	35399
D	3361	CCTTACACATTCCACGCTCCTGTTTGTGTTTTCCTCAACACGGAGCTACTCTTCGACTG	34200
O	3540	tctatggagcgcgcaaaagaagagcgcgtctcttcttgccgcagcgcgtcccaacagatctctgaag	35999
D	3421	GCCAGGGGAGGGGCTGTGTCACACGGTTGTGTGGGACAGCGTATGAGGGTAATTCCTTAAT	34800
O	3600	ccgactctacgcaccacgcagcagagcgcctccacgcgcgcgcgtcaagtlccacgcgtccag	36599
D	3481	CGCATCTCTCCCTACCGATGAAAGACGGGCTACATGTGTCACAAAGTTGCTTGAAAGTGGTT	35400
O	3660	ggaatatgagagagcgcgttaagtaactcttgtatataactgaagtgtgtgtgcgcggaaatatgc	37199
D	3541	GGTGATACTAAGACCGTCGATTTATTTTCCACTT-----GTCAACGACGGGGAACACAGT	35940
O	3720	gggcagccttatatacgcgcagcttctccacagacaaatcagaacatcaagatacgaatc	37799
D	3595	GGCAGCTGTATGCTGAAGCAACTCTTCCCGAGAGCATTAAGAAATTTGAGAGTACGATA	36540
O	3780	cgaggcaccgagcgcgcgcacgcagcgcgcgcgcgccttcttgaccgcgcgtcaagtlctcggtg	38399
D	3655	CGTGTCTACTGACGGTGGGACAGAACCAACGCGCTCTCCACGAGATGTGACCTTCAAGTGTG	37140
O	3840	gtctctgtaccacacagagggcgcgcacgaagtcttcagcgaaatagcagctactgtctgcctc	38999
D	3715	GTGTTCTGTGCCATPACAGGAGAACTTATCTTCCCTACGTGCACCTACACACTGTGTTCT	37740
O	3900	tctcgaagggttaaagaagccctccgcctgagagtttttgactctgcgcgaacgcagaagacttaaa	39599
D	3775	ATTGAAAGAAAGACCGGGCTTTTGGAACTCATATGAGCTGTGCGCGTGTGAGAGATCGGAAA	38340
O	3960	aaccacctctgagaagatgacttgcacaaatatactactcaagttatattagcgcgaacac	40199
D	3835	AACCATTTTGGCAGTATATGATTTGTTCACAAATATCTCTACAGAGATGTTGAGGAAATAT	38940


```

EEVYTRATDGTGEPRLSTDCFSVFPVPIQGEPIPTSTHTVALEKEAGLIERHE
LPRAEDRNHLCDSDCHNIYRIVDGNNDQPHFVDEITNVELEKEDRSSEYTLT
IAASNPSTGIALSTITITVYNREADOPFVFDITVAGISTSDSINRELLITIOATH
SENNPIYITIDLSMTVMDPTLASVRETRFILNPHTCVITLTIOTPAASHHGFQOVA
TDPAGYSDRANVKIYILISTNRNVEFLFNVTLEQVQNDLTAQFSAFEMTCNIDOV
VPADASGVINMGITEVGHFIKRDNPVPAEITLNDMDWLLAIOSTLARTLVLR
DLFTDTSAPADGSAAYALAYALASALLALCLIVLFIIRFKLNRLRLEALIVKY
GSVOSGLNRVGIABGTNRKHAVEGSPNPIWNETIKAPDPDSMSDASNDLIGIDLPH
FGENNPPRDYDERKTDKPEDIVATHNNNFKSTPSPERANOPKL"

misc_feature
4725..4803
/ gene="btr175b"
/ note="putative
transmembrane region"
primer_bind
5155..5172
/ gene="btr175b"
/ note="PCR primer site"
BASE COUNT 1482 a 1295 c 1224 g 1171 t
ORIGIN
Query Match 35.6%; Score 1959.2; DB 3; Length 5172;
Best Local Similarity 63.0%; Pred. No. 0;
Matches 3197; Conservative 0; Mismatches 1833; Indels 48; Gaps 9;
Oy 144 aactaaggatcgtgcaaaatgggggtgagaggtctctcccgaggtgtactgtctct 203
Db 7 AACAAAGGACATGTCAGAGTGGAGTTGACGTTGCAATTTCTAGTACACTCTTCTAATC 66
Oy 204 tttagctctgcgcgaactgccaacacgaagtlcgtacatcatcgaatcgaacgaagc 263
Db 67 TATGCAAGAACTGTATGTGGCTCAGAAAGATCGGCTTCATGTTGCCATACCAGACCA 126
Oy 264 gagactccggaactgcgcctcttgaatcgaaggaataatcatgagtggaacgcctta 323
Db 127 CCAGGCGGAGTCTCGTGAAGTACATTTGAAAGTCAAACTCGTCAAGTCAAGACACTA 186
Oy 324 ataccgcgcgcgaacccgaggaagatgatgtatgagaaact-----tctaccgat 374
Db 187 ATACCGGGGCGGACGACCGTGAAGATGTCTGATGACGGCTACACCGATGACACCTTACC 246
Oy 375 caaatgaticagtcatalatcatgaggaagaatacgaaggaagtcacatcgtcgaag 434
Db 247 TACGGAACCTCAGATATTATTACATGGAAGAAGATAGAAGAGAGATACCCATCGCTAA 306
Oy 435 cttaactatcaagggtccaacagc-----ggtgctgtcgatatagtcaggccagcc 488
Db 307 CTCATTTACCGGCGTCCAAATGTACATATGAGCCAGCATTTCTGTGTGTTCTTTC 366
Oy 489 agagccagcgtgagccctgaggtttcgacagaatgaagcagagcgcaatgagccttgt 548
Db 367 AATCTACTGTCGCCGTTATCCGACAGATACCAGATAGTAATGGAGAAATGGACCTTATT 426
Oy 549 attacgcaaaagcaagactacagacagcaacacatgacagactatgtlctcaatcaa 608
Db 427 ATTACTGACAGACAGATATTACGAGACTCCGGGAATGACAGATATGTTTCAACATCAG 486
Oy 609 gtgaggggtgatacacaagccgtactgtggtgcgtgagatagatcaatcgaagaacat 668
Db 487 ATATGATGGGGAACATTTGTGCGGGGTCTTTCACATGCTGATTGTAATAATGACCAAC 546
Oy 669 ccggccatccctgcaagtgctcagcgccgtcgtlaattccagaacatgagggagctagct 728
Db 547 GCACCTTATATTCAGACACTTCACCGTGTCAAGTTGACGAACCTAGGGGAAGCTGTCTG 606
Oy 729 accgagctcgtgtacaaagtgcagaccggaacggtgaatcagaacacgcgtctcaga 788
Db 607 ACTGATGCTGTATAGTGTAACTCCGATGCCCATCAGCAGCACCAGTTCATGACAG 666
Oy 789 ttccgtgtcgaagcagcagcggtgtcgaatgaagcatcttctaagtgttggaataaac 848
Db 667 TTTAGAGATCGAGAGATCG--AGGAGACGACCAAGATCTCTTACATCCAAAGGGCGAAC 723
Oy 849 gaccacagcagctgttcaatatgaagatgactgtgaggaatcaatccgtctgaactc 908

```

```

Db 724 ATTCGGGGGAGATGATCGGATGACATGACACGTTGGCATTAATGAACCGTTGAATTTC 783
Oy 909 gagacaactcagctcatatattatgagctacagctctctgactgcgtcaccgaacaac 968
Db 784 GAGACTAATCCACTACATATATTTCTGTGACCGCTTGGAACTCTCTCCAAACACTGAC 843
Oy 969 acggtcaccatgatgtgtgcaagtggaagcgtagagtgctggccccctgcgtgggtgag 1028
Db 844 ACTGTAAAGCTTATGTGTCAAGTATGAGATGTGGAACATCGGCCGCCACGATGGGTAGAG 903
Oy 1029 atcttccagtcagcagcttgacgaagaagactaataagagcttccctccgcgcgata 1088
Db 904 ATCTTGTGCTGTACACATTCGATATGAGAAACTGCTCAGAGCTTCCCTGTAGAGCTATC 963
Oy 1089 gacggggaacaggaatcaataatagggccatcaactataccctcaatcaggaatgacgtgac 1148
Db 964 GACGGAGATACAGCGCATCAACAAACCTATCCACTACAGACTGGAAGAAAGAGAGAGAT 1023
Oy 1149 gactctctt-----tccctgaggtgatgaaagcggagctatctcgaagtgc 1199
Db 1024 ACGTTTTCACATTGCGACCATCGAGGGGGGCGCTGACCGGCGCATATTATACGTAGAT 1083
Oy 1200 gagatcgaccgcaagacttgaaagagagctttcaacctaccatcgltgttacaacaa 1259
Db 1084 CCAATAGACAGGAGACACTTACAAAGAGAGAGTTCATTAATTCATATACCGCTATATA 1143
Oy 1260 tctactgcgttagcttgcacaagagggccacatlttcatcatcgtcaacgagctcaat 1319
Db 1144 TACGACAGATCTTTCAGCTACGCTACGCGCAACGTTGCATTAATGATCAACGATTAAC 1203
Oy 1320 gatcgcgaccgcgcgtcgtcataaagaatacagtatgtatatcatgtgaggaactcca 1379
Db 1204 GACCAAAACAGGAACCTTGTCTTCAAGATATCGTTTGAACATTAATGAGAGACACT 1263
Oy 1380 atgacttcaactcaatgaagaatlttgatccaatgacagatlttgagaaagcgt 1439
Db 1284 TTGACGTTGAACTTGATCGATCAGGAATTCGAGATTCATGATAGAGATCTGGGTCAAAACGCA 1323
Oy 1440 caatacagltggaacttgagagacgtgttccgcgaagggcggtgccatcttcatatc 1499
Db 1324 CAATACACCGTAGAGCTTAGAGAGCATATTATCCAGACAGCGGGCGCAAAAGCCCTTTATATA 1383
Oy 1500 ggcgcgggagcgctcaccagaagcagacctcatcatatgaggaacataaacacacatg 1559
Db 1384 GCACCCGAAGTGTGGCTACAGAGCAGACCTTCATCAATGAGGAGACGCTTAACCATTAATG 1443
Oy 1560 ctgagttacgaagatgltcatlttttcaagaacatcatcatlaaggttcaaaagcagtgtag 1619
Db 1444 CTCGATTTATGAAGTTCAGAAATTTTCAACGGATCAGGTTCCGGGTATATAGCANAGACATG 1503
Oy 1620 aacaacgctlaacacagtggtgcgagcgctgtgtlactgctgaacactgataactggaacgac 1679
Db 1504 GACATTAAGAAACAGCTGGGGTTGCTGATGTGATCATCATCATCTTAACATGGAATGAT 1563
Oy 1680 gaacttccatcttcogaaagagcagctaccgccgttcgtttaaaggaaacgctgcggc 1739
Db 1564 GAAAGCCTATTATTTTGAACATATGCGGTACAAAACGTTCTTCAAGGAAATGAAAGGAAG 1623
Oy 1740 ggtcccccgtggtgcacaggtgtcgtccctcgacagagaacatcagacagtgtagtgc 1799
Db 1624 GGATTTTTCGTGGGTGAATAGTAAAGAGCCACAGACAGACATTAATGACAGATGTAACAC 1683
Oy 1800 tcatgatgtggaacagcgtgtgtgaactcgtlcatagatgaatcaacggaagagatcttc 1859
Db 1684 ACATTAATAGGGAAGAACGCAAAATTAATATCTCTGATAGCAAGACACAGGGGAGATTCAT 1743
Oy 1860 gtgagcatgagcagatgcttgcagaccacggaagaaacactcatgtgtcagagtggc 1919
Db 1744 GTCACGCAAGAGCATCTTGTGACTATACCGACAAACGAACTTTTGTTCAGAGTCA 1803
Oy 1920 gctgacgatactttgggagagcgccacacaacacagtgacacacccagctgtgtagaa 1979

```

Dh	1804	CTGACGACACACCTGGGGGA---ACATTCCACACGGGCGACTTTCACCTACTGATCCAT	1860
Qy	1980	ctgaggaatgtcaacaacacactccctccacactacgttgcgcggttgcactccaagcgtc	2039
Dh	1861	CTGGACGATATTAAACAACACCCCTCCTACTCTGAGATTGCGCGGGAGACTTAACTGAT	1920
Qy	2040	gaggaagaacgttcccgaaagatatcgsatatcccggaattcaatcgtctaacgaccgagac	2099
Dh	1921	GAAGAGAAATGTCCCGGAGGGCTACATCATCACAAAGGAATCGAGCCACCGATCTCAG	1980
Qy	2100	accgcgcctacactgttgattcgcagatcgacttggagatcccaacttggccaccaagaagc	2159
Dh	1981	ACTGACGCTGAACTCGCGCTTCGAAATTCGACTGGACACACTTCTTACGCAACCAACG	2040
Qy	2160	agagaagaccaacctactgatacgtctgggtgtatgtatctgaaacgatatacccacc	2219
Dh	2041	AGGGAACTATCTCTATTGTGAATTTCCAAATTGTGTCGATTCGAAACAATATTCTGCG	2100
Qy	2220	gaaggcaaccggagttccgcactcggcgccctcgtgtgtgcaagagatccggagaacgtc	2279
Dh	2101	ATCAACAACCGTGGCTCGGCTATAGAGACGCTGTCGTAAAAAATACAGGAGACGTC	2160
Qy	2280	accactgacttcgaggaattcgcagatgcttaactccacccgtccggttgaagaccctac	2339
Dh	2161	ACATTGACTACGAAGAGTTCGAGATGTTGTATCTGACGCTGAGAGTCCGGACTTTAC	2220
Qy	2340	actgtcalcggaaatgtactcgaatgagcgacgttccagatcacataatacgcagtgaac	2399
Dh	2221	ACTGTATTGGGGATGTGATTTGATGATGAATCAACCTTCACCATCAATCATGATGATGAC	2280
Qy	2400	gacaaagcgccgactcctcgcgaaagcgaagctgtgaagcagaatgcgttgcgtagctg	2459
Dh	2281	GATTAATCCTCCGATATGCGTCCCGGGAACATGTGACACTCCCTCGTTCGGGAATG	2340
Qy	2460	gcgcgcgaagcgaacgactcgcgtcgtcgttcgcacacggaatctgaaagccgcgtctac	2519
Dh	2341	TCACAACCGCGGTGCTGATTTGGGACTCTGACGGCCACTGACATGTAGACGACACTTAC	2400
Qy	2520	aaccaagttgcgtctacactatatacaacctagaacaacactcccgagaggttaagtgaagt	2579
Dh	2401	AATCAAGGGGATACACTATTGAAGGCAACGAAGGCACACAGAGATCTATTGATGATC	2460
Qy	2580	gacttcaacaatgttcaaatltgaagtgtgagtgcgaacgaagcgatcogatgcagaacacc	2639
Dh	2461	GACTTCTACACTGTGTCGATCAGACAGTGAACAAAGCGGGCAATGTGATGATTCTCA	2520
Qy	2640	tggcgcttcaactgttaactaacacgcgtcactgcgtcgaagcagaagtgtccctgtgaaacgc	2699
Dh	2521	CGGAGATACAACTATATCTACACTGTGGTGCCACTGATGGTCTGACGCTGAAAGACCA	2580
Qy	2700	accgaatgctccccaagatlttcaactacttcogaagttccagcgcgatcogaatagaatc	2759
Dh	2581	GAGCATTTGCCCTGACGACCAACTTTTGTGGAAACATCCCGGCGAAGTGTGATTAACATC	2640
Qy	2760	atcgacaacaacaacaagtgtcctgscgcgtcactgtgaagtltcaacagacggtgtac	2819
Dh	2641	ATAGATACCAACAACAATAACTCTCAACCTGAAGAACGCAACCAATTTAAGCGGTCGTGAT	2700
Qy	2820	gtctgggagaatgcacagacgagcgagcagaaggtgttcagctgtactcccaagacggttac	2879
Dh	2701	ATTTACAGAAAGTGTGTGTGTGGAGATGAGGTGTGTAAGATTATTGGTAGCATTGGAT	2760
Qy	2880	agagaagagtgttaaccacacggtacagatacagaatgaacttgcggtgtaaaccccgagt	2939
Dh	2761	AGAATGTATTATTCACACACAATTCGTTACCAANTCACTACGCGGTGATCTCTGTTG	2820
Qy	2940	cgagattcttcgaggtgcacgtgcagacactggtgcgcttgaaggttcattaccgcggagac	2999
Dh	2821	CGGACATTTCTTCGCTGTGATTCGGATTAAGGGAAGAGTGTAGCTGTACTACACACCGAC	2880
Qy	3000	gaaaaattgacgcgcatgtggagtgcgctacaacatactatcttgtgaaatttcagat	3059
Dh	2881	GAGGTCTCTCGACCGGAGCGGAGCAACTCAACACAGAGATTTCTTTAACTCACTTCAT	2940

OY	3060	aactctttctctatggtgacggttagaagaaacccgagccgaagttgaaatatttgctt	3111
Db	2941	AACCTTTTCCAGCAAGAGACGGCAACACAAACCAAAATGATGCCGAAGTACTGGTCTA	3000
OY	3120	ctatgtgatgtgaacgacacgcctccctgtaagtaccatttccctgaagaaacctcggtttgat	3179
Db	3001	CTTCTTACGCGAAGCAACGACAAATGCTTCCAGAGCTTCCGAAGCCGAGTAACCTGCGTGGTGC	3060
OY	3180	gtttccgaagaggaagttgctgtgtgtccgtgtactccccaatcttaagcccggaag	3239
Db	3061	GTGCTCAAAAGTTTAACATAAGGGTAGCGGCTTCCAGCCACACATATTAGCACTGATATGA	3120
OY	3240	gataacacagacacccgacacactcgcgtgttcgggttaagaaatcttgagactacatcac	3299
Db	3121	GAGAACCTGCACACGGACACACTCTCGGGTCCGGGTACGCCATCATACGCTCACGATGCC	3180
OY	3300	gaaccgagacatcgagtggtccggatctctccacatgatcttgatgaaataacaaacttgg	3359
Db	3181	AACGGGGAATTCGAATGCCAGACGGCTTTTCACCATGATATCAATATCAAAAGTCACAGA	3240
OY	3360	gaacttagacccgtatagacttgtagggggtatctgggcacttaagaaatctacatg	3419
Db	3241	GAACTGGAAACCCCATGAGCTTGGAAGGCTATTGGGGAACTTAATGCTATTCAATTAA	3300
OY	3420	gccttcgacacaggtctaacccgcagcagaggttccaagagacgtlatacccttgcataccg	3479
Db	3301	GCTACGACCCAGGACATCCCGCAACAGATGTCCAAATGAATACGACCTCGTGAATTGCT	3360
OY	3480	ccctacaactctcaacacccctgtgtcgtgttcctccgaaccccgactccgttalcctgctt	3539
Db	3361	CTTTACACTTCCACGCTCTGTGTTTTGTGTTCCTCAACACGAGCTACTCTTCTTGACTG	3420
OY	3540	tctatgagagcgcgcaaaagagagcgagcgttctctgcagcgcctgcacaagatctcttgag	3599
Db	3421	GCCAGGGAGAGGCGTGTCTCAACGGTTTGCTGGCAAGCGTAGAGGTGAATTCCTTAAT	3480
OY	3600	ccgaatctaacgccacgcagagagagcgctccacgcggcgagtcagttccacgttccag	3659
Db	3481	CGCATCTCGCTACCGATGTGAAGACGGGCTACATAGCTGGCAAGTTGCTTCMAATGCTT	3540
OY	3660	ggaatagagagcgcttctcaactttgatataactgtagtggagcagggagaaatgac	3719
Db	3541	GGTATACTGAAGCCGTCGATTATTTCOCATT-----GTCAAGACGGGAGAACAGT	3594
OY	3720	ggagcagcttatatacgccagcttttccagagcaaatcagaaatcgaagatcagatc	3779
Db	3595	GGCACTCGATGCTGAAGCAACTCTTCCCGAGACATATAGAAATTTGAGGTGACGATA	3654
OY	3780	cgggcacgcgacgcgcgcgaagagccgcgcgcgtttgagccgaagatcaggttttcggt	3839
Db	3655	CGTCTTACTGACGGTGGACAGAACCAAGACCGCTTCCACGAGACGCACTTCACTGATG	3714
OY	3840	gtctcttaccacaacagggcgagccagtttcaagcgaaatgtagcttctgtcctt	3899
Db	3715	GTGTTCGCGCGAATCAAGAGAACCTATTCTTCCATGCTCATCACTCATCTGCTTTC	3774
OY	3900	ttcgaaggttgaagaagcctccgttgtaggtttttagctgtgcgcgaagcaagaacctaa	3959
Db	3775	ATTGAAAAAGAAAGCGGCGCTTTTGGAAAGTCATAGACCTCCCGTGTGAGATGGGAA	3834
OY	3960	aaccactcttgcgaaggttgacttcgaagatattactaaggttttatgtgaaggcaaac	4019
Db	3835	AACCATCTTTCGACTGTGATTTCTCAATATCTAATAGGATAGTTGATGGAAATTAAT	3894
OY	4020	gagggtctttctgactcagacagtcacgaacaaagtatctcccttgcgcagaggttgcac	4079
Db	3895	GAGGCGCACTTTGGACATCGACAAATTAAGAACGTTCTATTCTCGTSTAAGGAATTTGAT	3954
OY	4080	cgcgaggtggcacgcttctaacgcgtgcacaatccgcgcgcgcgcgaactcgcgcgaagcaact	4139
Db	3955	CGAGGTGTACGCGAAAGTATTACATTGACGATAGCTGCCAGCAACTGCAC-----CGCA	4008

[illegible]

RESULT	7	LOCUS	AR096316	5577 bp	DNA	PAT	08-SEP-2000
ACCESSION	AR096316	Sequence 1 from patent US 6007981.					
VERSION	AR096316	GI:10025017					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 5577)						
AUTHORS	Bulla, L.A.						
TITLE	Polynucleotide encoding a receptor for a Bacillus thuringiensis toxin and methods of use						
JOURNAL	Patent: US 6007981-A 1 28-DEC-1999;						
FEATURES	Location/Qualifiers						
source	1..5577						
BASE COUNT	1476 a 1496 c 1404 g 1201 t						
ORIGIN	/organism="unknown"						
Query Match	Best Local Similarity	35.6%; Score 1954.6; DB 6; Length 5577;					
Matches 3248; Conservative	0; Mismatches 1769; Indels 62; Gaps 17;						
Qy	152	atcatgcaaaatgggggttgagaaggtctctccagcaagtgctactgctctcttagcctc	211				
Db	187	ATAATGTAGATGGCAGTTGACGTCGGAATGCGTCCTTCGGTGTATATAGCGCC	246				
Qy	212	tgcgcgaactagccaaccaagatgtctgtaactatcgcaataccaagaacggagactcc	271				
Db	247	TGCAGTTTAACTCTCAAGAGAGATGTGGGTATATGACCCCAATCCCAAGCTACCAACGACC	306				
Qy	272	ggaaac---tgccgcctatgtattagcaagaaatcatgaggtgagaaacgctctaatacc	328				
Db	307	GGATATATTTGGCAGTACTTAATATTTGAAAGGCCAGACATGAGATCAGAGGCCCTGCTCC	366				
Qy	329	cggcccgaccggagggagagatgatatgagaact-----tcttacggatccaat	379				
Db	367	CGCCCGGAGGGGATGACCTGTGCATGGACGCGTACACACGATGAAACAGCAACCTCGG	426				
Qy	380	gattcaggtcatatcatatgagaggaagaaatcgaaaggaacgltcatcatctcgaaactaa	439				
Db	427	CACGAGGTCATCTCTCAATGATGAGAGATGAAAGAACAAATACACATCGCCATCTTAA	486				
Qy	440	ctatcaaggtgtccaacaacagccggtgtgtgtgattatgtccaagccagcccaagcct	499				
Db	487	TTAT-AACGGACATCAACTCCGTTCAATGAATGCCATTTTATCGGTTGATCAATC	545				
Qy	500	gggcccctaagtttcagacaagaatgaagccagacggccaatggagccctgttattacgaag	559				
Db	546	TGCTGATCCCGGTATCTACGAGAGATTTGACAAAGGAGTGCATC--TCAATCATACGCAAG	603				
Qy	560	acaagaactacgagacagcaaccaatgacagatgtgttcttcaatccaaagtggaggtga	619				
Db	604	ACAGCATTAACAGATGTGCCCGGCATCACAAGTACTACTGTTCATGTGCGCGTGAGAGCGCA	663				
Qy	620	atcaacagccgctactggtgtgtgtgtgtgtatagttaactatgaacatcgacgaacaatccgccatctc	679				
Db	664	GTCGTGTGTGGCAGGCGGTCTCTCTGATATGCTCAACATTAATATGCAAGGCGCCATAT	723				
Qy	680	gcaagtgtgcagcgcttcgtaatttcacagaacaatggagagctaaactgaacgaactgct	739				
Db	724	ACAAACTTTCAGACGCTTCCGGGTTCTCTGAACCTGGGCGCAGCAGGCTTGAACAGATGAC	783				
Qy	740	gtaccaaagtgtcagaccgagcaggttgaatataagaacccgcgtctcatgacgttccgtgtcga	799				
Db	784	ATACCAAGTATTCGAGCGCGGAGCGAGCGAGATCAGCACAAGATTCAATGAGCTTCAGAGATGA	843				
Qy	800	cagcagcagggcgtcagatgtgaagatctcttcatatggttggagaaatacgaaccccaagcga	859				
Db	844	CAGGTTTCG---TGCGCGAGGAGACCTTTTACATGCAAGGAGCAATATATCCCAACCA	900				

QY	860	ctgtctcaatatgaagaagacgtgtggtgatacttcgccttcgaacttcgagaacaactca	919
Db	901	ATGGATGTGGCTAAATATGATCCATAGGGCGTTATATACCTGCTCAACTCTGTGTACCCAGTCC	960
QY	920	gctcatatattgaagctgaacagcttcctgaactcgctacagcaacaacaagcttcaact	979
Db	961	GCTGCATATATTACAGCTGACAGCCCTGGACTCGCTCCGGACACCCCAACGGTGTACTAT	1020
QY	980	gatgtgtgaagtgtgaacgtagagctctcgccccctcgctgtgtgtgagatcttcacat	1039
Db	1021	GATGTGCAAGTGGGGAATGTGAACAGCCGTCCGGCGGCTGGCTGGAGATCTTCCCTGT	1080
QY	1040	gcacagacttgaagagaagaactcaatcagaactctccctccgcgcgatagaagggagac	1099
Db	1081	CCMAACAGTTTGAAGAGAAATCTTACCMAAACCTTACACAGTGAAGGGCATCGACGAGAAC	1140
QY	1100	gggaatcaatagggcatcaactatatacctctcagggatgaagcgcaagactcttttc	1159
Db	1141	TGAGATCATATATGCTTATCATCACTACAGGCTGATCACAAATGAGGAAGAACATCTTTCAG	1200
Q	160	ctgtgaagtgtat-----gaagacgagactatctgtgcgctgtacgtgaagctcagc	1210
Db	1201	CATTGAGGGCCCTGCGTGGTGAAGAAAAGGGGGCTGTATTCTCGTTCGGCAATTGACCG	1260
QY	1211	cgacaagcttgaagaagagacttccaacctcaaccatcgltgtcttacaatctactyagc	1270
Db	1261	CGACACACTGCAACGAGAGGTTCCTCACTTACGATGCTGCTTACAAATATGATATGAGA	1320
QY	1271	tagcttgcgaagaagggcccacatttcatctatcgtcaacgaagctgaatgaatgaagacc	1330
Db	1321	GGCTTCTCCATCATCAACAAAGGTGCTCATCTGTGACAGACATATACGACCAAAAGACC	1380
QY	1331	cgagccgctgcataaagaatacagatattgatalcatgtgagaaacccaatgacttaaa	1390
Db	1381	TGAACCTTATACAAAGGAATATCGACTGGCAATFCAATGAGAGAGACCCCTGTACCCGTCA	1440
QY	1391	cttcaatgaagaatttgatctccatgagatctgggtgaacgcgtccaatacaagt	1450
Db	1441	CTTTCATTAAGAAATTCGGATTCATCATATTAAGATTTTGAAGTTCACAAAGCCATGATCA	1500
QY	1451	ggaacttgaagacgtgtctccgcgaagggcgcgctccgacttaccatcgccgcggag	1510
Db	1501	GGCTTAGAGACGCTGGACCCCTCCAGGGCCTGCTGAGGCAATTTACATAGCCGCTTAAGT	1560
QY	1511	cggtaccagagcgagaccttcatcatgtggtgacacataaaaccaacatgtctgattaca	1570
Db	1561	CGGTAACGACGACAGACTTCATCATGCGCACCCCTCAATCACTCATGCTGATTAACGA	1620
Q	571	agatgtcaattttcagaatacatcatcttaagtccaagcagctgtgaatgaacaaagctag	1630
Db	1621	ACTGCCAGATTTGAGAGATTAATACGATTCGGGTGTGAGACCGAACACACACACAGAG	1680
QY	1631	ccaagctgtggcgagcgctgtgtacgtgaacctgatacctaactgaaagcaagaattcccat	1690
Db	1681	GCACGTGGGGCGTGGCGTTGGTTCACATTTGACCTCATCATTTGGAACAGATGAGAGCGGAT	1740
QY	1691	cttcgaagagagcaactactccgcgtctgtttaaggaagacgttcgcgcgcgtctccggt	1750
Db	1741	CTTCCAAACGCGCGTGGCAACACCGTCACTTCGACGAGACTGGAAGGGGAGGCTTCTTCGT	1800
QY	1751	ggcgaagctgtccctctgacagagagacatgacagcagctagtagtgatctatgtatgg	1810
Db	1801	CGCCAAAGGGGTGTGACACAGACAGACATCGGGGATGTGTGACATATCTTTATTGGG	1860
QY	1811	caacgctgtgtactaccctgtcatagatgaatcaacgagagagatcttcgtgagcatg	1870
Db	1861	TACCGGTGTTAATTCTCTCAATCATGACATGACAAACACACCGGGGACATCCGGCTCTCAGCTA	1920
QY	1871	cgatgccttgcataccacgagaaacactcatttgttcagggtgcgcgtgcagatac	1930
Db	1921	CGACTCTTCAACTACCATTCAGAAAGTGAATTTATTTGTGACAGGTGCGAGCTACAGACAC	1980

Db	3058	GGTGGTGGACCGGTGATGTGATGACCAACGACACCGTATCTTTTAAACCTCATGACAA	3117
Oy	3062	ctctcttcttgatgttgtagaggttagagagaaccagagacgaagtgtgaataatcttcgcttc	3121
Db	3118	CTTTCATGGGGGAGAGAGAGTATACAGAAATACAGAAACGACACGAGTCTCGTTATCT	3177
Oy	3122	atgtgagtgtgaacgaacaacgctctctgtgatgtcatctgtcgtatgaaacctcggttgatgt	3181
Db	3178	GTTGGATGTCGATGATGACATGATCGCTCGTAATTTGCCACCGCCAGGACGATCTCTTGGACAT	3237
Oy	3182	ttccgaagagcagttgcgtgtgtgtccgttgcgttaccctccagaatctaacgacggagaagga	3241
Db	3238	ATCTGAGACCTTAAAGCAGGGCGTCGCTCTTGAACCAACATATTTCTGCCCGGACCGCA	3297
Oy	3242	tgaaccaaagacggagaacaactcgtctcgtgtttagcgaatctctgacacctcaagatcacga	3301
Db	3298	CGAGCCCGACACAGACAACTCCAGGTCGGCTACGAGATCTGTAACCTC---AGCACGA	3355
Oy	3302	ccgagacatcgaagtgccggaatctctcacatgatctcgatgtgaacaaacaacctgysga	3361
Db	3355	GGGGGACATCCAGAGCGCGAGCTGTTTGATGATACAGATGCGCAACGTCACGGGAGA	3414
Oy	3362	acttgaagccgctctgtgactgaggggtatttgggcacttcgcaatcttcattcgtgaagc	3422
Db	3415	GCTGGAGACCGCCATGAGCTCAAGGAGATTTGGGGAGCTAGCTATACATACGGGC	3474
Oy	3422	cttcgacacagcgttaccgc	3481
Db	3475	ATTGCAACACGGCATTTCCGCAATATGTCATGAMCAGACATATAGACTGATCATCATCC	3534
Oy	3482	ctacaacttccacaaccctgtgtctgtctgtctccgcgaaccgcgactccglatctcgcttc	3541
Db	3535	GTTCAACTACTACCGCTGCTGTTCTCTTCCGACCAAGATGCGCATATCGACTTGC	3594
Oy	3542	taggaagcgcgcgaacagagaagcgctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	3601
Db	3555	GAGGACAGACCTGTATTCATGAGATCTTAGCGACAGTGAACGAGATTTTGGACG	3654
Oy	3602	gattcaacgc	3661
Db	3655	GATATCGCGCATGTATCCGACGGACATCCACGGGGCGTCTACCTTCCAGTGGTAG	3714
Oy	3662	aaatgaggaagcgctctcaagtattgtatataacttgaagtgtgagcagagaaataagcgc	3721
Db	3715	CGATGAGGAATCACAAAGGATCTTCAAGTAGTAAACGATGGC-----GAGAACCTCGG	3768
Oy	3722	gcagcttatataagcgaagcttttccagagcgaatccagacaattcagaatcagaatccgc	3781
Db	3769	CTCGTTGAGTTTACTGTGAAACCGCTTCCAGAGGATACAGGAGTTTCCGATTAACGATTGC	3828
Oy	3782	ggccacgaagc	3841
Db	3829	CGCTACAGACCAAGGAAACGACCCAGGACCGCTGTCCACGGACATGACCTTACAGTTGT	3888
Oy	3842	ctcgtaccacaacagagcgaaccagtgatcagcgcgaanaatgtagctactgttcgcttc	3901
Db	3889	TTTTGTGCGCCGCAAGGAAACCTAGATTGCGGTCTCAGAAACATGCTGTGCTTCAT	3948
Oy	3902	cgaaggggtgaagaagcctcgtgtgagagtatttgagcttgcgcgcgaagcagaagacctaa	3961
Db	3949	AGAAAAGAGTCCCGCATGTGAAAGATCTACCAACTTCTCTAGACACAAGATCAGAA	4008
Oy	3962	ccaactctgcgaagaatgacgcgaagatatctactacaagtattgttcgcgcgaanaagga	4021
Db	4009	CCATCTCTGTGAACACATCTGACACGATTTATCTATCTATTTACGATGGGACAGGA	4068
Oy	4022	gggtcttcttgtaactggaacgcgaacgaacgctcatctcccttgcgcgcgaagtgtgacgc	4081
Db	4069	AGGTCAATTTGGGCGCTGATCTCTTGGCAACAGGTTGTTCTTGAGAAAGACGTGATTAAG	4128
Oy	4082	cgaagtgagcagcttcaacgcttgcacatctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	4141

Db	4129	GGAAACAAGTGGCTTCCCACTACTGTGCAAGTGGGGCTAGTAACTGCCCGATG---	GTGG	4189
Qy	4142	gattccctctgagagattccatcctctgtgtcaccggtcaatgtaagaagacgacccg		4201
Db	4186	CAATTCACCTTCCCTGGTTCCATTCCTTACTGTCACGTGTTACCGTAGAGGAGGACACCTCG		4245
Qy	4202	cccaatttcgagcaagcactttaacagcgggcaatttcgaagtttggaagcatgtgcg		4261
Db	4246	TCCACTGTTTGTGAGGGAAATGTACACCGGAGGAAATACACAGCGCATCTCATGGGAG		4305
Qy	4262	ggaattgtttactctgcaaggcgagccacacagaagacgacacatacagttacacattat		4321
Db	4306	AGAGCTGCTTCAGATTACATGCGACCCAGTGTGAAGCTGGCCATTACTTATGCTATAGA		4365
Qy	4322	ccgtcgagacatgtcagcttggacagcaagcacttagaagccgtgcgcgactcgcgcgt		4381
Db	4366	CTACGATACATGGTAGTGAAGGCCACCGCTGAGGCGAGTAGACAGTGGGCTTTGTACT		4425
Qy	4382	gcattgagaccgcggcgtgtcttcgtctcaatatgcaagcccaacgcttccatgcgcat		4441
Db	4426	GAACCTCAAAACCGAGTGTGTGAGCGCTTATATCCAGCCACCGCCAGATGATCAAGACT		4485
Qy	4442	gttcgaattgacgtctactgccttagatgaagttctgtaacatcgacaaagcccggtgaa		4501
Db	4486	GTTTCAATTTCAGAGTACAGCTATGTACAGCGCGCGGCTCAAGACCGCACAGTGCAC		4545
Qy	4502	agttaccctcatctcatctcgcaaaacccgcgtgacattcatatttcgataaaccaacttgag		4561
Db	4546	CGTGAAGTGGTATTCGTGGGAACCGCGTCTACTTCTGTTTCGTACACAGCTGCACAA		4605
Qy	4562	cgtttagcagaagcaaaattcattagcgcgcacgtttcaagacccgggttcaacatgacgt		4621
Db	4606	GGTCACAAACAACAGAACTTTATCGGGGACACTTACAGCGTGGGTTTCATCATCACTG		4665
Qy	4622	caacatcgacaagttgtgtccgtttcaagca---cagcaagcgcggttggcgcaagaagac		4678
Db	4666	CAACTTCGACCAAGTGGTGGCCGCTTACAGACCCCGTACACCGGGGTGGCGTGGACACAG		4725
Qy	4679	caccgaagttgcgcgcgcgaactlcatccgggacaacgctgcgcgttgcagggcaagaagttcga		4738
Db	4726	CACGCAAGATG-CCGGCCACTTTCATACGGGACAAACGTACCCGTAATCCGTGATGAGATAGA		4784
Qy	4739	ggccgtccgcagcgacaaggtgtcgtctgcgacacatccagctgatcgttgcgacaacaag		4798
Db	4785	-CAGATCCGAGTACCTTACCTCTCGACTTCGATACAAACAAACAGCTG-CCGGCGCGAT		4842
Qy	4799	ccgttgtctgtaagacctgttgaacgggtgaacatccgaagcgttagcgagggtgtaatgca		4858
Db	4843	CGTGGTGTTCGAGGACTTGTGTTGACCACTTCACACCCGCGACTTCGGCGCTGCATCGAAC-		4901
Qy	4859	gattcgccctactaagcactaagcgcgcctcgcgtgtgcagcttccctcgtcgttact		4918
Db	4902	-CTGCACGGTACGTCTGGCCTCACTGCTGTGTGTGCTCGGTTTTCATGTGCTGTGTCT		4960
Qy	4919	gcttctgcgcatlgttctgttagaacaagagcacttgaacccgycagctgcgaagcacttccat		4978
Db	4961	ACTGCTTACTTTCATTCATCAGAGACTAGAGGGCTTAACGACGAGGTGTGAAGCCCTGTGCAT		5020
Qy	4979	gacgaagtaagcgtccggttgaatcccggtctgaacccgcgcgggtc---ggccggcgac		5035
Db	5021	GACGAAGTACGGCTCACTGTGACTCTGGAATTGAACCGCGCCGCAATCGCCGCCCGGAC		5080
Qy	5036	caacaagaacagccgtctgaaggctcaaaccccaatgtgaaacggaggtccatccgcgcgca		5095
Db	5081	CAACAAACACACTGTGGAAGGCTTCCAACTTCTTCATGTGAAGCAATTAAGACGCCAGA		5140
Qy	5096	cttcgaagcatagtgacgcgagttgagcacttcgcgacttcggtcatcgtacgtgaagcatgc		5155
Db	5141	TTTGAATGCCATTTAAGSAGGTTCCAAAGCACTCTATGTGATCGGCATCGAAGATCTTGG		5200
Qy	5156	gcaattccgcgaagcacttctccgcgcggcgacacaga		5194
Db	5201	GCACTTTGGCAAGCTTTCATGTGATCTTAGGTGAACA		5239

RESULT 8
 LOCUS 177078 5577 bp DNA PAT 03-APR-1998
 DEFINITION Sequence 1 from patent US 5693491.
 ACCESSION 177078
 VERSION 177078.1 GI:3013232
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 5577)
 AUTHORS Bulla, L.A. and J.L.T.
 TITLE Receptor for a Bacillus thuringiensis toxin
 JOURNAL Patent: US 5693491-A 1 02-DEC-1997;
 FEATURES
 source Location/Qualifiers
 1..5577
 BASE COUNT 1476 a 1496 c 1404 g 1201 t
 Query Match 35.6%; Score 1954.6; DB 6; Length 5577;
 Best Local Similarity 63.9%; Pred. No. 0;
 Matches 3248; Conservative 0; Mismatches 1769; Indels 62; Gaps 17;

844 CAGCGTTCG---TGCGGACGAGGAGACCTTCTACATCGAAGCGAGCAATATCCCAACCA 900
 QY 860 ctggttcaatatagaagatgactggtggaatcgaatcgccctgacttgagacaactca 919
 Db 901 ATGAGATGGCTTAATATATGACATAGCGGTATATACCTGCTCAACTTGTCACCGACTCC 960
 QY 920 gcttcataatattaagagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 979
 Db 961 GCTGCATATATTCAGGCTACAGCCCTGAGCTCGCTCCGGAACACCCACGCGTACTAT 1020
 QY 980 gatgtgtcaagtgtggaagacgttagagcttcggcccccctcgctggtgtggaatcttcag 1039
 Db 1021 GATGGTGCAAGTGCGCAAAATGTAACAGCCGCTCGCCCTGAGATCTTGCTCT 1080
 QY 1040 gaagagcttgagaggaagaaactaatcagagcttcctcccgcgagatagagggagacc 1099
 Db 1081 CCAACAGTTTGAAGAAATCTTACCAAAACTTCACAGTGAAGGGGAGTGAAGGAGACAC 1140
 QY 1100 ggaatcaataagggccataactataccctcctcctcctcctcctcctcctcctcctcctc 1159
 Db 1141 TGAGATCAATATGCCATATGACATAGAGCTGATCAACAAATGAGGAAGACATTTCTTACG 1200
 QY 1160 cctbgaagtgatt-----gaagacggagctatctcgaagtgactgagacgcg 1210
 Db 1201 CATTGAGGCCCTGCGCTGCGGAAAAAGCGGCGCTGATCTCTGTCGCCAATTGACCG 1260
 QY 1211 ggaacagcttgaagagagcttctcaactcaccatcgcttgcttgaatctcaactgagcg 1270
 Db 1261 CCAACACTGTCAACGAGAGAGGTCTTCCACTTACGATCGCTTCAAAATATGATGAGGA 1320
 QY 1271 tagcttgaacaagagccacatttcaatcactcgtcgaacgcgtcgaatgatacagcagc 1330
 Db 1321 GGCGCTTCCACATCAACAAAGGTGATCATTTGACAGACATCAACAGCAAAAGACC 1380
 QY 1331 cgaagccgtgcataaagaatacaagatgataatcagaggaagaaactcgaactctaa 1390
 Db 1381 TGAACCTATACAAAGAAATATGACTGGCAATCATGAGGAGAGCGCCCTGACCTCA 1440
 QY 1391 ctccaatgaagaatttggatctccatgacagatggttggatgaagaagcctcaatacagc 1450
 Db 1441 CTTTCGTAAGAATTCGATTTTCATGATTAAGATTTAGTTCAAAGCGTCACTAGACGCT 1500
 QY 1451 ggaacttgaagacgtgttcccgccaagggcggtccgacttcaatccgacgagggag 1510
 Db 1501 GCGTTAGAGAGCGGTGACCCCTCAAGCGCTGAGGACTTCTCATATGAGCCCTGAAGT 1560
 QY 1511 cggctaccaaagacgacttcaatcagaggaacataaacaacacacatgctgattaga 1570
 Db 1561 CGGCTACCAAGCAGACACCTTCATCATATGAGGCAACCTCAATCACTCCATCTGATTA 1620
 QY 1571 agatgcatctttcagaacatcatccttaagttcaagcagagtgagacatgaagaacgtag 1630
 Db 1621 AGTGCCAGAGTTTCAAGATTAATTCGATTCGGGTGTAAGGACCGACACACAGACACGAG 1680
 QY 1631 caacgtgagcgaagcgtgtgtgtagctgaacatgaaacttgaacagcaacttccat 1690
 Db 1681 GACAGTGGGCGCGGTGGTTGACATTGACCTCATCAATTTGGAACAGAGAGAGCGCAT 1740
 QY 1691 ctctgaagagagacgactactccgctgctgttaagagagacgctcgcgcttccggt 1750
 Db 1741 CTTTCGAAACAGCGCGTGCACACCGCTCACTTCGACAGAGACTGGAAGGAGGAGGTTTCT 1800
 QY 1751 ggcacaggtgctcgccctcagacagagaaactgaacagcgttagtgaatctatgtaggg 1810
 Db 1801 CGCCAGGCGGTTGACACAGACAGAGACATCCGGGATGTCTGAGACATCTTTATTGGG 1860
 QY 1811 caacgttgtgactactcgttcatagatgaatcaacagggagagactcttcgtgagcatgga 1870
 Db 1861 TTAAGCTGTTAATCTTCACCATGACAAACTCAACCGCGACATCCGGCTCAAGCTTA 1920
 QY 1871 cgatgcttctgactacacagagagaaactctatttgttctgagtggtggtcgagagatcc 1930
 Db 1921 CGACTCTTCAACTCACTCACTGAGAAAGTGAATATTGTGAGGAGTGCAGCTACAGACAC 1980

[illegible]

QY	3002	aaaatlgagccgcgatggtgggtatgagctctcaacatcatcttctgttaattcatcgataa	3003
QY	3058	GGTGTGGAGCGGTGATGGTATGAAACCAACGACACCGATATTTCTTCAACCTCATACAA	3117
QY	3062	ctctcttctcgtatgctgacgttaagagaacaccagaagatcttgaattctgtctct	3122
Db	3118	CTTTCATGGGGGAGAGAGAGGTTAACAGAAATCAGAACACACAGAAAGTTCTGTTATCTT	3177
QY	3122	attgtagtltgaacagcaaacgccttcctgatatgtcaattgcctgtatgaatcccggtltgatgt	3183
Db	3178	GTGTGATGTGAAATGACAAATGCTCTGTAATTTGCCACCGCAGGAGAACTCTCTTGACATAT	3233
QY	3182	ttccgaaaggagcagttctgtgtgtcgtgtactcccaagaattcatcgaccgcgagacggga	3244
Db	3238	ATCTGAGAACTTTAAGCAGGGCGTCCGCTCTTGAAACCAATATCTTCCCGACCGACGGA	3297
QY	3242	tgaaccagagaccgagacaactcgcgtgtccggtttagacgatctctgactctaagataccga	3303
Db	3298	CGAGCCCGACACACAGCAACTCCAGGGTCCGCTACGAGATCTTAACTC---AGCACGGA	3353
QY	3302	ccgagagacatcgaggttgcggaatactcttcacacatgatctcgattgaaacaaactctggga	3363
Db	3355	GGGGGACATTCGAAGTCCGGAGCGTGTGTGATGATATACATATGCCAAGTCACGGGAGA	3413
QY	3362	acttgaagaccgatgatgacttgaaggggtatgtgggcaacttgcgaataatltcaattgaagc	3422
Db	3415	GCTGAGAGCCCGCATATGGACCTCAAGGGATATTGGGGACGTACGCTATCATATACGGGC	3474
QY	3422	cttgcagacaggtctaacccgagagaggtccaaagagagtlacaacctgttatccggc	3483
Db	3475	ATTGCACACAGGCAATTCGGCAAAATGTCATGAAACGAAACATATGACCTATCATATCC	3533
QY	3482	ctacaacttccacaaccctcgtgtcgtgtctccgacaaccgactcgcgtatccgtcttc	3544
Db	3535	GTTCACATCTACTACCGCGCTGAGTTGCTGTTCCGACCAACGATGCCGTATACGACTTCC	3599
QY	3542	taggagagcgcgaacagagaagcgcgcttctgcgcagcgtctgccacgagttccctggagcc	3603
Db	3595	GAGGAGACGACTGTAAATCAATGGAGTGTACGACAGTGAACGAGAGATTTCTTGAGAGG	3655
QY	3602	gattcaagccaccgcgagagagagccctccacgcgcgggcagcgttaagttccacgtccagag	3666
Db	3655	GATATCGCGGCACTATCCGACCGGACACACGGGGCGCTGTCACCTTCAAGTGGTAGG	3714
QY	3662	aaatgagagagccgttcaagtaacttgaataacttgaagtggagagcagaggaataatagcgg	3722
Db	3715	CGATAGGAATCACACAGGTRACTTTCAAGTAGTAAAGATGCG-----GAACAACCTCGG	3766
QY	3722	gcagcttatatacgcagacttttcccaagagcaatcatgacaatltcaagatacagatccg	3783
Db	3769	CTCGTTGAGGTTACTTGGCAAGCGGCTTCCAGAGGAGATCAGGAGATTCGGATTAACGATTCCG	3822
QY	3782	ggccacgcgagcgcgcgacagagagccgcgcgcgttggaccgcagcgttaagtttccgttgt	3844
Db	3829	CGTACACACACGAGGAACGACCGACGACCGCTGTCCACGAGATATCAGCTTACAGATGTT	3888
QY	3842	cttgcgtacccaacagaggcgacaccagtgcttgcagcgaaaatgacgcgtctgcgtctctt	3903
Db	3889	TTTTTGTGCCACGCAAGAGATCTAGATTCCGGCTCTCAGAAATGCTGTGTGGCTTTTAT	3948
QY	3902	cgaggtgtgaagaaagcctcgcgtgagagtttttgagctgtcgcgcgaagcagagaaccttaaaa	3963
Db	3949	AGAAAAGAGTCCCGGCAATGGAGAGTCTACACAAATCTCTTATGACACAGACATCAGAA	4008
QY	3962	ccaacctgtcgagaatgactgtccaagatatctactacaggtttatattgacgcacaacaga	4023
Db	4009	CCATCTCTGTGAACACACATGTCACAGCAATTTACTATCTGATTTATTCGATGCGACACGGA	4068
QY	4022	gggtcttttcttgaatgacagtcacagcaagcagctcatctctcttgcgagggagattgacgc	4081
Db	4069	AGGTCATTTTCCGGCTGGATCTCTGTTGCCAACAGAGTTGTTCTCTGAAGAAAGAGCTGATTAAG	4128
QY	4082	cgaggtgtgcacgcttctaacgcgttgcacatcgcgcgcgcgaacactgcgcgagccactgtg	4141

Db 4129 GGAACAAGTGGCTCCACACTGTCAGATGGCGGCTAGTAATCTGCCCGGATG---GTGG 4185
OY 4142 gatccctctgcaagcttccatctcgtgtgtaacatgtaagaagcaaccgcgc 4201
Db 4186 CATTCACCTTCCCTGGTTCACATCTTACTGCTACCTGTTACCGGAGGAGGACCCCTCG 4245
OY 4202 cccattcttcgagcagccttaccagcggcatttcgaagcttgaggagagattggccg 4261
Db 4246 TCCAGTGTGTTGGAGGAGATTTACACCGGAGGATATCCACAGGAGATCCATGCGCAG 4305
OY 4262 ggaattgtactgtcagggcgagccacagaagacgacaccatcaagctacacataga 4321
Db 4306 AGAGCTGCTCAATATACATGACAGCCAGCTCAAGCTGCGCATTTACTTATGATATAGA 4365
OY 4322 ccgtgcgagcagcagcgtcggaagcagccttagaagccgtgagcagcttgccctgcgc 4381
Db 4366 CTACATATCAATGTAGTGGAGCCCGACCTGGAGGAGTGAACAGTGGCTTCTGTACT 4425
OY 4382 gcatgcgacccagcgcgtgtcttcgcataatgacagccacgcattccatgacgcgc 4441
Db 4426 GAACGCTCAACCCGAGTGTCTACGCTTATATCCAGCCACGCGCAGATGCTGATGACT 4485
OY 4442 gtctgagctgcagctcatcgctacgagataagcttctgcaatcgacacacccgtgtgaa 4501
Db 4486 GTTCAAAATTCGAAGTCAACAGTACTGACACGCGCGCTCAGACCCGACCGAGCTGAC 4545
OY 4502 agtcaaccctcatcgacaaacgcgcgtgaccttcatttcgataaccactgtgagac 4561
Db 4546 CCGTGTACGTGGTATCCTCCAGAACCGCGCTCTACTTGGTTCGTACACGCTGCACACA 4605
OY 4562 cgttagcagaacagaatattcatagcgcacagcttcagcaccgcggttcaacatgctg 4621
Db 4606 GGTGCAAGACAAACAGAGACTTATGCGGACACCTTCACGCGTGGGTTCAACATGACCTG 4665
OY 4622 caaatcgacacagctggtgctgctcagcga---cagcagcgcgtgtgacgaagacac 4678
Db 4666 CAACATCGACCAAGTGTGTCGCCGCTACACCCCGTACACGCGCGTGGCCTGAGACAG 4725
OY 4679 caaccagctgagcgcagcttcatcgagaaacagctgcccgtgacagcagaagagctga 4738
Db 4726 CACGCAAGATG-GCGGCCACTTATACGGGACACGATCCGCTGCTGATGAGATATA 4784
OY 4739 ggcctccgcagcagcagcagctgctgctgcgacacacacagctgagctgtagacacaag 4798
Db 4785 -CAGATCGGATGACACAGTCTCTCTGAGCTCGATACAAACAAAGCGTG-GCGGGCGCAT 4842
OY 4799 cctggtctgcagaagcctggtgagcgggtgagcaactccgagcgtagcgagagagtcagca 4858
Db 4843 CCGTGTGTTGGAGGACTTGTGACCAACTCCAGCCCGGACTTCGGCGCTGACTCGAGC- 4901
OY 4859 gatcgccgttcaagcagcagcagcgttcgcgtgctgtaggcttcctctgcctgact 4918
Db 4902 -CTGCAACGGTGAAGCTCTGGCTCTCACTGTCTGCTGTGCTCTTCATGCTGCTGTGCT 4960
OY 4919 gcttctgcattgtcttgaagacaagacagcagcagcagcagcagcagcagcagcagcagc 4978
Db 4961 ACTGGTTACCTTCAATCAGCAGACTAGACCGCTAAACGAGCGGTGGAAGCCCTGTGCT 5020
OY 4979 gacgaagtaagcgtcgtgagcagcagcagcagcagcagcagcagcagcagcagcagc 5035
Db 5021 GACGAGTGAAGGCTCACTGAGCTCTGAGTTGAACCGCGCGCATCGCCGCCGCGCAC 5080
OY 5036 caacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5095
Db 5081 CAACAAACACATGTTGGAGGCTCCAAACCCATATCTTCAATGAAGAAATTAAGACGCCACA 5140
OY 5096 ctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5155
Db 5141 TTTAATGTCATTTACGAGGCTTCCACAGACTCTGATCTGATCGGCATCGAAGATCTTGC 5200
OY 5156 gcaattccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5194

Db 5201 GCACCTTGGCAACGCTCTTCATGATGATCTGAGGTGAACGA 5239

RESULT 9

AF367362

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

10..5208

/note="s1 allele; CryIac-susceptible"

/codon-start=1

/product="cadherin-like protein"

/protein_id="AAK8196.1"

/db_xref="GI:15149240"

/translation="MAVDVRLTAAVILAAHLTVADDCSYMAVAPRPERDPENQNF

EGVPMQNPILPADREDVCAAPDPAFVAVTIFPMEELGEGVALAARLRYCTNP

TVPVPLNFGFHLGPIYRIPEOGGMHLVITQROXYENPMQOYFENRVDEPGE

ATVPLIYNIDNNAPITOMPERCDIPREGGCTECKYVSDAGESTPMPTQITS

DRNDEEYFELVRNIGOMKTVHRLINKLDTLEENLHLEFATLADLPNVTVM

WVQENIESRPMMEIFAVQDFEKTAAQAEFVADIDGIDIPRIETESRD

LEFVETIGAGREGAMFKAYIDRDTLEEFVHSLIAKYGDNDVSGSPSESTDIY

IIVNDVDAVPPRPSYIEIMEEAAMTLLEDFGEGFDRGLGHAQYVTLISISA

GAHEAPYLAPEVGOOSFIVGTQNHHLDEVEFQKIQRAVAIDMDPRVAGI

INILINMNDLPIFEHDVQVTEKTEGAFVATYLAARDIIDDVHSHLMGNAY

YLSIDKDTGDLVITIDAFNHRONELFVQIRADDTIGEPNTNTAQLVLODINT

PTLRLEPTTPSVEENPDGFIPTLEHADPDPTAELRESIDMTSTATQGDADA

EEFVNCIEIEFVYENLNDRTAIGRVVREIREVTLIDYEMFVLYLVRVTDLNTVI

GDDYDISFETIIEIDMNDPPLWEGTLOEFREVAASGVIGSVLATIDGPLN

QVRYTIPRLTPEPDIVEIDNSGQISKKRQAIDADBPPOHLYTVVASDKDLS

VDVCPDPNPNYNTGDTIHTTQNNYVPRAEKPEFIVYVGAARDVHVDLPS

DLDDEYTHKVSQINAINRPLRDOFEVDLETLVYVYNNAGKLRDDEPHRIE

FNVIDNEFGSDGNRNOETQVLLVLDINDINPELPLGSLWDISEGLDQVRYTPI

FAPDRDEGDNSSAAVDIVLSPTDRDITLPQLFTMTLEKDGIDQGTLEAFMDL

RGWGTATYHKAVDHPORISYEKPLVLRPNFDPVFPPOPGDTILAKERAY

VNGVLAITYDGEFLRIYATDEGDLHAGVFSISGDDEALQFPVFDNGVLAITTE

QLPEDPREPVYTRANDGCTEPGPRSDCTITVYVPTGSEPFVSTYVAIERD

AGMEERATLPLAKDPNRMCDCHDYISTVGSNGHFAVDQSNELFTLLERA

BOEHHTLILIGASDPSPAVILQASTLTVYVWVRANRPVQSLAYAGISTDPTNR

GLTLHNTSHSEGLPVYTLVODSMEDASTLOAVDETAFNLNPOGVLTLPQAPSHH

GMPEFDWATVTGSETARTKVKYVLSIRNVEFTFPMNLTLEVEPNDDTAETPLF

GMRCNITIOALPASDPATGAARDQETVRANHIRDDLPVPAEVLCEVLTGPNPTAT

ALQENLNLADLTFGETPITLGEQAKRAVIALAVALAALCVLLILFETPITRAN

RRLALSMTKYSQDSGLNRYGLAAPTGNKRAVGSNPINNETLKADFPALSSQSD

SGLIGIEDLPQFRNDYFPPDESSMRGVYVNHMGANSVANHHNNFNFATPFSPBEA

NSOARR"

BASE COUNT

ORIGIN

Query Match

1393 a 1433 c 1417 g 1112 t
35.4%, Score 1946; DB 3; Length 5355;

Best Local Similarity 63.3%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 1790; Indels 93; Gaps 14;

QY 228 caacagatgttcacatattatcgcaataccaagacggagactccggaactgcgcgctatt 287
DB 76 CAAGATTGTTCTTAACAGTAGGCAATTCCTACAGCAACAGAGGACCGAGATTTTCTTAATCAAA 135
QY 288 gattacgaagaataatcatgagatgaaacagacccttaataccgcccagaccgagagaa 347
DB 136 AATTTTGAAGAGATGATGATGAGATGAGAACCCCTGTTACAGCGAGGATPAGGAGAT 195
QY 348 gtatgtatggaagactcttaccggaatcaatga-----tcaagttacatacatagag 401
DB 196 GTGTGATGAACGGGTTGCGACCCGAGTGCTTGAAACCCCTCAACCGTCATCTTCATGAG 255
QY 402 gaagaaatcgaaagagaagcgtcatcaatgagaagcttaactatacaaggtcaaacagcg 461
DB 256 GAGAGAGATCGAAGGGGACGTGGCCATTGCCAGGCTTAACTAACCGAGATCAATACCTCG 315
QY 462 gtgctgcgattatgtcagcgccagccagcgccagctgagccctgagtttgcacaga-- 519
DB 316 ACCGTGTAACTCCATTAACTTTGGTACCTTCCACTTGTGGGCGGTCATACGTAAG 375
QY 520 ----atgaagcagacagcgcaatgagccttgttattacgcaagaacaagctacagaca 575
DB 376 ATCCCGAGCAAGGGGGGAGCTGGCATCTTGTTATTAACGAGAGGAGACTATGAGACC 435
QY 576 gcaacacatgacagactatgtgtctcaatcaagtgagagggtgaatcaacagcgactg 635
DB 436 CCCAACAATCGACAGAGATATCTTCAACGTGAGGGTGAAGAACAGGCCCGAGAAAGCACT 495
QY 636 gtggcgtggaagatgatacatcgacagaaatccgccaactcttgaagtgtgtcagcgcc 635
DB 496 GTATGCTCATCTATGTTCACATTTGACGACAAAGCTCTATCAAGATGTTTCGAGCCT 555
QY 696 ttggtatctccaagaacatgagcaggttagacatgacagactgcgttacaaggtgtaaac 755
DB 556 TGTGACATTTCTTGACACGGCGGAAAGGGGACACACAGAAATGCAAGTACGTGGTACGAT 615
QY 756 cgcgacggtgaataacagcaacgcgctcaatgaagttccgtgtgcagacagcagcggtgca 815
DB 616 GCTGACGGGAGATGACACACGTTTCATGACGTTCCAAATCGAGACGATCG---AAAC 672
QY 816 gatgaagaacatcttaccatggtgtgagaaatagacccaacgacgtgttcaatlatgaag 875
DB 673 GATGAAGATATTTTCCAACTCGTCAAGAGAAATATTCAGGGACGTGATGCTGCAT 732
QY 876 atgactgttggtgaataatcgcccttgaactctgacaacaactcagctcatatattacg 935
DB 733 ATGAGGCTTATCTCAAAACACCTGTGACTATGAGAAACCCGCTGCATTTGTTAGA 792
QY 936 gtacagctcttgaactcgttaacgaaacaacagcgttaacatgattgtcaagtgtgag 995
DB 793 GTTACAGCTTTGGATTCCCTACCAAAAGCTTCAACAGTAAAGATGATGGTGCAGTAA 852
QY 996 aagctgaagatcgcgcgcctcgtgtgtgagatcttctcagtgagaggtttgacgag 1055
DB 853 AACATGAGAGCAGACACACAGTGGATGAGATCTTGCCGTGACGATTCGATGAGAG 912
QY 1056 aagactaatcagaagcttccctccgcgcgcatagacgagggaacggaatacaatagagc 1115
DB 913 AAGACAGCAACAAGCTTTCAGGGGTTGAGGCATCGATGAGACACGGGAATTCGATAA 972
QY 1116 atcaactatacctcatcaggtatgacgctgacactctt-----ccctg 1163
DB 973 AATTTTATTAAGATTCGAAACTGAGAAAGCGAAGAAAGTTTGTACGTGTGAAACATA 1032
QY 1164 gaagttatgaagaagcagactatctgcacgtgagatgagatcgacgagcaagcttga 1223
DB 1033 GAGAGCTGTGCGGAAGGTGCTTGTTAAAGTGTGCTCAATAGACAGGACACTCTAGAA 1092
QY 1224 agagagctttcaacctcaacatcggtgtgttacaacttactagacgtgac----- 1274
DB 1274

DB 1093 AAGGAATTTTCCACGTCGTCTTAATAGCGTACAAATATGGGACAAATGACGTGAAGA 1152
QY 1275 -----tttgcacagagggcccaatlttcatcatgttcaagagctcaatgttcag 1325
DB 1153 AGTCCGTCAATTCAGATCGAAGACGATATGCTATTATGTGAACACAGCTGATGTACAG 1212
QY 1326 cgaaccgagccgctgtcataagaatacaglatgtatcatatcagtaggagaaactccaagact 1385
DB 1213 GCCCCGCTGCTATCCGCTCTTATCTACATTGAAATTTATGAGGAAGCTCGATGATA 1272
QY 1386 ctaaacticaatgaagaatttgattccatlgatcagatltgggtgaaacgctcaatc 1445
DB 1273 TTGAATTT---AAGAGACTTTTGTTTCCAGATAGAGGTCTTGTCGCGACGACAGTAC 1329
QY 1446 acagttgaaacttgaagagcgtgttccgcgaagggcgcgctccgacttcatatcgcgcg 1505
DB 1330 ACAGTCCACTGTGAGAGCATCTCCCGGGGGGCGCCAGAGGCGCTTCACTACGCGCGC 1389
QY 1506 gggagcggtcacagaaggagacactcatcatatgaggacataacacacacatgtcgat 1565
DB 1390 GAGTGGGCTACCGACGACAGTCTTCATCGTGGCACCGAAGCAATCATGCTGCAC 1449
QY 1566 taagaagatgtcatlttcaagaacatcatcatlaaggtcaaaagctgagatgaacaac 1625
DB 1450 TTTGAAGTCCAGAGATTCCAAAAGATACAACTTAGGGCGGTACCCCTACACATGGACAT 1509
QY 1626 gctagccaagttgagcagagcgctgtgtacgtgaaccgtatcaacttgaacgaact 1685
DB 1510 CCCAGGTGGTTGGTATPCCGATTAATAACATTAACTGATCACTGAAACGATGAGCTG 1569
QY 1686 cccatcttggagagagagcagctactcgcgtgtttaaaggagacgctggcgcgcttc 1745
DB 1570 CCGATTTGAGACGACAGCTGACGACTGTCCCTTCAAGAGAGGAGGCGCTGGCTTC 1629
QY 1746 ccggtgccaaggtgtctgcgccttcagacagagacatcgacagctagtgatcatcatg 1805
DB 1630 CCGGTGCGCACGTCTTGCGCAAAAGGACAGGATTTGATGATAGATGCAACATCTCTCA 1689
QY 1806 atgggcaacgctgttgaactactgttcatagatgaatcaacgaggagagatcttcgtgagc 1865
DB 1690 ATGGGCAACGACAGTGAACCTTGAATATGACAAAGAACCCGGGACATCTCGTGACA 1749
QY 1866 atgagagatgcttgcataccacagcagagaacactctatgttgatcagtgcgctgac 1925
DB 1750 ATCGAGATGATTTATTAATACACAGACGAAGAGAGCTTTGTGTAGATGAGAGCTGAT 1809
QY 1926 gatactttgggcgaagcccaacaacacagtgacccaacgaactgtgtatagaactgag 1985
DB 1810 GACAGTTGGAGA---GCGTATATATACCAACACCGCCCACTGGTATACAGCTGCAA 1866
QY 1986 gatgtcaacaacactcctccacactgaagcttgcgcgcttgcattccaagcgttcgaagag 2045
DB 1867 GACATCAACACACACCTTCACTGACAGCTGACAGCTCCGACAGCACTCCGATGGAGAG 1926
QY 2046 aaggttcccgaaagatagagatatcccggaagaatcaactgttcaaccgacccggaacagc 2105
DB 1927 AACGTGCCGAGGGGTTCTGTATCCACCGAGCTGCACGCCACGACCCGACACACAC 1986
QY 2106 gctactcgtgtgtcgaagatcgacttgaactcaccctggtgccaaccaagcagggcagagag 2165
DB 1987 GCCGAGCTCGCTTACGATGATGAGTGGGACATTCCTATAGCCACCAAGAGGAGAGAT 2046
QY 2166 accaacctactgataagcgtcggtgtatagttatcgaagaagatataccccaacgagggc 2225
DB 2047 GCTGATGCTGAGAGATTGTTATGTAAGAAATCGAGACGATATACCCGAATCTTGAA 2106
QY 2226 aaccggggttccgcacatcgcgctcgtgtgtgcaagagatccggggaacagctcaacatc 2285
DB 2107 GACCGAGGACCGCATCGCGCGCTGTGTGCTTCGAGATCCGGGAACACGTCACCAATC 2166
QY 2286 gacttcaggaattcgaagatgtcttaccatccagctccggttgaaggagactcaaacatgct 2345
DB 2167 GACTACGAGATGTTCGAGGTGCTGTACTCACCGTACAGGCTACGAGATCTGAACACGGTC 2226

QY	1755	tgctgcgccctcgacagagaaacatcgcagcagtgtagtgcattcattgatlgtggaacgcgtg	1818
DB	1781	tcagaggctccacgaccgggacattagacacacgaccttgcatttccattcttgggaaattgcga	1840
QY	1819	ttgactactgtgtcaatgaatgaatcaacccgggagagttcttcgtgagcatgagccgtgtct	1878
DB	1841	atnacttttttgaagaaatccacgaamaaacttgccacatnactgactattgtgtgacggct	1900
QY	1879	tcgactaccacacgaagaacacttatlttgcagtgccgctgtcgaataacttggcg	1938
DB	1901	tcgatttatccacagacagaaatgaatttaactacaaatgtgcgcgtcagacacacatgtcgg	1966
QY	1939	a-----cgcgccacacacacacagltgacccaccagctgtgtagaaactgtggaatgtcaaca	1995
DB	1961	agccagactccacagccatnaccagccgcttcacgttcagctgacatgacactgcagacac	2020
QY	1996	aaacccctcccccaccctcagctgtgcgcgcgttcagcttccaagcttgagagaaacgttccg	2055
DB	2021	acnacccttcttacttctgagcctgtccttgccgttaagctcgtctgtgaaagaaatgtgtccag	2080
QY	2056	aaggaatacagagataccccggaaatcaactgtcactccgaccgcggaacacacgctacgt	2115
DB	2081	agggctttgaaatcacaacccggagaaataccgccacacgaccctgacacnccagatnactgc	2140
QY	2116	ggttcgagatctgacttggactccacactgtggccacaaagcaaggcgagagagaccaccta	2175
DB	2141	agtttgaattgattttgggacacatctcttccactaaacagggcgctgattaccattccaa	2200
QY	2176	ctgaaatacgtcgtgtgtatgttatcgaaacgatatacccccacgagggcaacccgggtt	2235
DB	2201	tagaacttccacggatggcgatataaaccatcttccaaacccagccgacacacgaag	2260
QY	2236	cgcacatcgcgcgctcgtgtgtgcaagagataccgggaaacgctacacatctgcactcgag	2295
DB	2261	agcgctgtggggccagatggtgacacaaaggagatccgccatnactgtacattcttggag	2320
QY	2296	aattcgagatgtcttacctcaacgcgtccgctgtgagggactccaactgtlcatcgagata	2355
DB	2321	agtttgaattttctctactccacacagtgagattcgggactttgcacacagatnaccgacag	2380
QY	2356	actacgatgagcgagcgttcaacgataccaataatcgacatgaagcaaacgcgcacatc	2415
DB	2381	attatagagaatttacccttccacatcgatataatataatnactgtgaacacaaactggcctatct	2440
QY	2416	tcgcgacgagcagcgttgcagcagagatgscgtgtgcagactgtgcgcagcgagcgacgc	2475
DB	2441	ggcgctgtgttttcttcttaccacacacaccttccatgatttccggacggatnacttaccggcgtgc	2500
QY	2476	tcactggtctccgtgtctgcgcacagacatcgcagcggccgcttcaacaaagtgcgtlaca	2535
DB	2501	tcactggggtccgactacgtacagacattgatggccactttacaaacaaagtcgggtaca	2560
QY	2536	ctatacaacctcagaacaacaacctcccgaggatagtgaagtgtacttcaacaactgtc	2595
DB	2561	ccatttatccccacgaagatnacttccgaagctcttactccagatatttctgattacgggctc	2620
QY	2596	aaattgaggtgtgactgcgaacgagcgatcatgacgaacgaaccccgcgcttcaactgt	2655
DB	2621	aaatttaccgttgatgagaaatggtgcactgacgcgtatatttccacctcgtttggcacctca	2680
QY	2656	actacacccgtatcgtctagcagcagtgctcccttggaanaacgcgaacggaatgtcctccag	2715
DB	2681	actnaccagcttatnaccagacaaatgtttcgaagaaatgaagaaagaaactgtccccgg	2740
QY	2716	attccacactactcgaaggttccacggcgatatcggaaatagaaatcatcgacacaaaca	2775
DB	2741	atccagctttcttggaatnacttccgggacacatncttaatttaccctctgcacataaaca	2800
QY	2776	aagtgctcggagccgtcaactcagagaagtltcaaacccgcgctgtgactcgtggagaatgca	2835
DB	2801	aggtcccgccagcagacacacacttcatgctgatttcaacaaacggctgtactttttatgaaatgtac	2860
QY	2836	cgaagcgcgcgaagagtggtlccagctgtlactccacacacgcttaccagagacagatgtacc	2895

Db	2861	CCGATTTCACGAACCTGGTCTCAAGATATATCCATCCATCGACGAAAGCAAGACGAAATATATAC	29220
Qy	2896	acacggtacacgaataacacgaatgaacttcgcgttgaacccccacgaactcgcggagttcttcgaag	2955
Db	2921	ACACGGTGGCGGTACCAAGATCATCAATTATGCTGTGAACACACGGCTGCGAGACTTCTTGCCA	2980
Qy	2956	tggacctgtgaacactgyltcgccttgtagtggatlaaccgggggagcgaaaaaatltgaacgcg	3015
Db	2981	TGAGACTTGGAATTCAGGGCAGAGGTATACGTGGAGAAAGAACCAATATGAGCTCTGGATCGGG	3040
Qy	3016	atgggagtgagcctacacatactactcttltgtaattcaatcgtataactcttctctatg	3075
Db	3041	ACAGAGGGGAGAGACCAACACACAGGATATTCATTAACCTCATTTGACCAACTTTTATAGGAAG	3100
Qy	3076	gtgaagcgttaggagaaacacaggaagaagtltgaatatltgtcgtcttatgtgagtlgagac	3135
Db	3101	GAGATGGAATATATGAATATATTAACACTCTACAGAGTGTCTGTGTATCTATTATGATGAAATG	3166
Qy	3136	acaaagcctccttgagatgcatcttgctcgtatgaactccggttgaatgtgttcccgagaaggacg	3195
Db	3161	ACAACGCTCTGAAATATGGCCGACTCCAGAGAGCGTGAAGTTGGACCATTTCCGAGGATTTAC	3220
Qy	3196	ttgctgtgtgtcc-----gtgtactccgaaatctacgcacccgagcaggatgtaac	3244
Db	3221	AAGAGGGTATTAACACTCGATGGCCAAAGCCATGTGATATACGCCACCGGATATATAGCAAG	3280
Qy	3247	caggaacggaacaaactcgcgtgttggttaggaatctcgtgaacctcaagatccacgcgcgag	3300
Db	3281	AGGACACGCCAAACCTCTACCTTGCTGATCCCAATCTCGGCATGACAGTCCACCAATATAG	3340
Qy	3307	acatacga---ggtgcgcgatatctctcaactgaatctcagatltgaacaaacaaactltggagac	3363
Db	3341	ACCTGGACACGTTCCTCCAGACTTCTCAACATGCTGTGCGCTATACCAACGTAACCGGATTC	3400
Qy	3364	ltbgaaccgctatgaacttgaaggggatlttgggacactltgggacactlaaatcatltgaagcct	3422
Db	3401	TCCACAGACGCAATGCTCTTGTAGAGAGATATTGGGGAGCTTACGATATATAGTATAGTGGCT	3460
Qy	3424	tgcgaaccgctctaccgcgacgagagtgctcaagagaagctacacccctgylcatccgcctc	3488
Db	3461	TGCACCAAGGGATATCTCCAGACGATATCTCATGAGGTATGTAATGGAAATTTCCACACTT	3520
Qy	3484	acaacttccacacccctgt	3543
Db	3521	ACAATTTACAATCTCTCTCAGTTCGTTTCTTCTTAATCCGGAGACGATTTCTACACACTGGCTT	3580
Qy	3544	gggagcgcgcaacagaaagcgcgcttctbvgcagcgcgctgcacaaagatctcctggaacga	3603
Db	3581	TGGAAACGGCGAGTGTAAATATATGTTTGTACCTTGTAACCGGTGACCCGTTAGACAGGA	3640
Qy	3604	tctaaagcacaacgaagagaagggcctcaacgcgggcagcgtcaactgttccacgtcccaaggaa	3666
Db	3641	TACAAGCAATTGACGACGATGTGCTTATCTGCTGGCTGTGATCTTCGATTTGTTGGAG	3700
Qy	3664	atgaggaagcgccttcagttacttgaatactgaatltgagtgagcagaagaaatagcgggc	3722
Db	3701	ATGCTGATGCTCAACTACTCTTGAGATTAATATATGCGCAGACCTTTGGAACCTTGT	3766
Qy	3724	agctataatlaagccagcttctccagagcaaaatcagacaattcagatcaagatccggg	3783
Db	3761	TGCTGACA-----CAGGCGCTTCTGAGGAAGCAAGGAATTTGAGTTTACATTCGCGG	3814
Qy	3784	ccaaagcagcgcgcaacgagagccggcccgcttggacgcgaagctcaagcttctgtgtgtct	3843
Db	3815	CTACAGAGGGGGAAGCAAGACTCTGATCATATTTCAACAACACTTCCACTATTAACAGTCTCT	3874
Qy	3844	tctgaaccacaacagaggaaccagtgltcaacgaaatagcagctactgtctgcgctcttcg	3903
Db	3875	TGCTTCCGACTTTTGGGTGATCCGATCTTTTCAAGATTAACACTTATCTCAGTACATTTGTTG	3934
Qy	3904	agggtgaagagggcctcgttgaagattlttgaagctgcgcaagcagaagacctaataacc	3963

[illegible][illegible]

Dp	488	CGGATCCGTCATTCGGCTCTCTAAGGAGCGGCACACAGAAAGCGGGGTCCTGGGACG	547
Oy	3580	ctgcacacagatctccctygaagcgcatctacgcacgcgaaggaagcgccctcaagcgggca	36389
Dp	548	CTGGCAAGAGACTTCCTGGAGCCCATCTACGCCACGCACAGAGCGCCCTCCACGGGGCA	607
Oy	3640	gcctcaagcttcacagctccaaagaaatagaaagcgcttaagactttatataactag	36399
Dp	608	GCGTCACGTTCCACGTCACAGGAAATAGAGAGCGCCTTCAGTACTTTGATATTAACCTAG	667
Oy	3700	tgggaagcagagaaataagcgggcagctataltatcagcagcttlctccagagaacaatca	37598
Dp	668	TGGGACGAGGAGAAATATAGCGGGCAGCTTATATTACGCCAGCTTTCCAGAGCAATCA	727
Oy	3760	gacaattcaagatcaagatccgggcacagcagcgcggcacggaagcccgccgcttgya	38191
Dp	728	GACAATTCAGGATACAGATCCGATCCGGGCCACAGACGGCGGCAGGAGCCGGCCGCTTGG	787
Oy	3820	ccgaagctcaagcttcttggtgtcttgtaaccaacaagggcgaccagtgctcagcgaa	38797
Dp	788	CCGAGCTCACGTTTTTCGGTGGCTTCTGTACCCACGAGGGCCACCAAGTTCACACGAA	847
Oy	3880	atgcagctactgctgcctctcttggaaggatgaagaagccttcgctagaagtttggcgc	39399
Dp	848	ATGCAGCTACTGTGTCCTTCTTGAGGGGTGAAMAAGCTTCATGAGATTTTGGCTGC	907
Oy	3940	cgcgaagcagaagaccttaaaaaacacctctcgaagaatctactgcgaagatctactca	39999
Dp	908	CGCAAGCGAAGACTTAAAAACCACTCTCGCAGAAATACCTGCCAAGTATCTACTACA	967
Oy	4000	ggttatttgcgcgaacaacagcaggctcttctgctactgacagctcaagcaagctacatc	40599
Dp	968	GGTTATATGACGGCAACACAGAGGGTCTGTCTGCTGGACACAGTCAAGCAACGTCATCT	10277
Oy	4060	cccttgcgaggaatttgaccgcggaaggtggcgaagcttaacagctgaacaatcgsggga	41119
Dp	1028	CCCTTGGCGAGAAATTTGGACCGGAGGGTGGCCACGCTTACACGCTGCAATCGGGGGA	10877
Oy	4120	gcaactgcgccgaagcccaactggagatccctctggaagcttcaatctctgttgcacgctca	41797
Dp	1088	GCAACTGCGCCGAGGCCACTGGGATCCCTGTGAGACTTCATCCTCGTTGTCAAGGTCA	11477
Oy	4180	atgtaagaagaagcgaacccgcgcacaaatttcgagcagaaccttaacagcgggcatctt	42399
Dp	1148	ATGTAAAGAGAAAGGAACCCCGCCCAATTTTGAGAGAGGACCTTTTACACGCGGCAATTT	12077
Oy	4240	cgaagcttggacagcatitggccgggaatttgctacgttcagctggcgcggaagagaagcg	42999
Dp	1208	CGAGCTTGGACAGCATTTGGCCGGGAATTCCTTACGTCACGGCGAGCCACACAGAAAGCG	12677
Oy	4300	aacacatcaagctacacacatlagacgctgcgaagatgcagcttgcagacagcagcttagaagcg	43599
Dp	1268	ACACATCATCGTATCATATGACCGTGGCAGCATGTGAGTGAACGACGCTAAGAGCGG	13277
Oy	4360	tgcgcgaactgcgccttcgcgtcgtatgcgaacacacgcgcgttcttcgtcctaataatgaac	44119
Dp	1328	TGCCGCACTCGGCTTTCACCTGCACTGCACACACACGCGCGCTTTCGCTCAATATGACG	13877
Oy	4420	ccacgcgcttccatgaagcagatgttttgatctgaagctatgcgtacagctacagctctg	44799
Dp	1388	CCACGCGTCCATGACAGGGATGTTGAGTTTGGACTCATCCCTTACGGATATACACATCTTG	14477
Oy	4480	caatcgcacacagccgctgtgaagcttacctatctcatalcgcgaacaaacgcgcgtgcactca	45399
Dp	1448	CAATCGACACAGCGTGGTGAAGTGTACTCTATCTCATCGAAACCGCGCTGTCTTCA	15077
Oy	4540	ttctcgataaccaacttgcagccgttgcagcagaacagaatttcaatagcgcgcgttca	45999
Dp	1508	TTTTTCGATTAACAACAACTGTGAGCCGTTGAGACACAAAGAAATTTCACTACGGCCACGTCA	15677
Oy	4600	gcaccggcttcaacatgcgttgaacaatgaccaggt 4636	
Dp	1568	GCACCGGGTTCAACATGACGTGTGAATGATGACCAAGT 1604	

[illegible]

Db	220	CAGGCGAGATGCATGGCTGCTACATGTGCTGAATACCCAGACGACAGACAGGCGTGCATCTTC	279
Oy	280	cgccattatgtatcgaagaataatcatgtatgtgaacagcttcataccggccgaccc	339
Db	280	CACCTCAAAATTTTGGACGGTTTAACTGGGCTCCAGCAGCACCATATTATACCAAGCTGAGATC	339
Oy	340	gaagagagatgtatgtagaaccttaccgatacaatga-----ttc	384
Db	340	GAGAAAGGCTCTGCTCTCAATGACTATGAACCTTGATCCCTGGAGCAACAACATGGTGACC	399
Oy	385	agtcatacatgtagaggaagaatctgaagagagagctcatcttgcaagcttaatac	444
Db	400	AGGAATATTACATGAGGAGGAGATCGAAGGTCCGTAGTCATTGGGAAAATTACTAC	459
Oy	445	aagggtccaacagcggtgtcgttcgatatgtcagggcagcccaagccca---gctgg	501
Db	460	AAGGAACACCCCTCCCAATATAGATTACCTTTTGTTGGTGGAGCCCAATGCTTG	519
Oy	502	gccttgatttcg---acagaatgaagcagcgccaatggagcctgtctatgaacaa	558
Db	520	GAGCAGAAATTCGGATATCTTGACCCACTGAGACTGTGATCTTGTAATTACTCANA	579
Oy	559	gacaagactcagagacagcaacatcagagctatgtcttcaatcacaagtgaggg	618
Db	580	GGCAGAGCTATGAACATCCCTGATATGAGAGATACACGTTCCATGTGATGTGGAAGGC	639
Oy	619	aatacagccgctactgtgtgtgcgttgagatagatcaaatcgaagcaatccgccatcc	678
Db	640	AGTCGCTGTGTGTAAACGCTGAGCTGGATATGTGAAATCAGCAGCAATGCGCCATCA	699
Oy	679	tgcgaatgtgtcagcgctgtgttaatccagaacatgtcgaggtcgtagctgagccgctcg	738
Db	700	TTGAGATGTTAGACCTTGCACTTACCCGAACTTTTGAACCCCATGTTTAAAGATGTA	759
Oy	739	tgtaaccaagtgtcagacccgagcgtgtgaatctgcagcccgctcatcaagcttcggtgc	798
Db	760	AATATATCGTGTCCGACGCGAGCGGTCTAGTACAAAGTGTATGATATATATATAG	819
Oy	799	acagcagcagggctgcagatgtaaagcatctctacatgtgttgagaatacagcccaagc	858
Db	820	ACACGCGAGAG---AGGAGACGAAAAAGTATTCACATCGATACGAAAAAGATTTCCGGCG	876
Oy	859	actgttcaatatgaatgtacgtgtgtgtgaatcttcgcaccttgaaacttcgagaaactc	918
Db	877	ATTGAGCAAGAGTATATATGTTCTTGAATTTGAAAAAATCTTGTGTTGCAAGAAATC	936
Oy	919	agcttcatatatgtatgcatacagcttcttgaactcgcgtcacagcaacaacacagatcaaca	978
Db	937	CTCTACACATATTAGAGTCAAGGCTTCTTATTTCTTACCAAAACATTAAGACCGTGCTCA	996
Oy	979	tgatgtgtcgaagtgtgaagcgtatagctgtgcgccccctgcgtgtgtgtgaatcttcag	1036
Db	997	TGATGTGTGAAGTGAAGACGTGGAACATATGAATAATCCGTGATGAGGAGATCTTGCTG	1056
Oy	1039	tgcagcagtttgacgagaaagactaatcagaagcttccctccgcgcgatatagaaggggaca	1099
Db	1057	TGCAACAGTGTGATGAAGAAAACAGCGCAAAATCGTTTCAACAGTGCACAGCTATTGATGGGACA	1116
Oy	1099	cgggaataaataaggscataactcaataccctcaacgaaggtatgcgtgtccgactctttt	1156
Db	1117	CGGGATCAATATAAACCTATATTTCTATCTGTATACAAACTGAAGATGAAGCAAAAGATTCT	1177
Oy	1159	-----cccttgaggtgtatgtgaagcagagactatctcgaacgctgactgagatcg	1200
Db	1177	TCAGCATTTGAGAACATATAGGGGAAGGCGACAGACGSGTGCGCAAGATTCACAGTGGCTCTATAG	1236
Oy	1207	accgcgacaagcttgaagaagagcctttcaacctcaacatcgctgc-----	1255
Db	1237	ACACGAGCTACTGAAAAAGGAGATATCTTATTAAGAAATATTCATATTAACAAGAGTG	1296
Oy	1253	-----ttcaaatctacgaagcgttagctcttgcaacagagggcccaacttcatcatcgta	1300
Db	1297	ATAATGAGAAAAAGAGTGAATCATCGTTTGAAGACTTCAGCAAAATGTGAGATTATATATTA	1356

OY	1309	acgcgctcaatgtctcggcgaccgccgctgcataaaagaaatgaagttagatcatcttg	1368
Db	1357	ACGATTAATAATGTATGCAGAGCCCAAGAACCCTTCCATTAAAGAAATACAGATTCATTAATGG	1416
OY	1369	aggaactccaatgacctctaactcttaactltaagaagaaatttggatccatbatcatgaaatttg	1428
Db	1417	AAGAATAATGGGATGACTTAAATTG---CAAGAGTTGGTTTCATGACCGTGACAATGG	1473
OY	1429	gtgaanaagcttcatacacacagfytgaacgttgaagacgttgttcgcgcaggcgctccg	1488
Db	1474	GTCGCCACGCTCAGTACGAGCTCATCTTAGAAGATACATACGCCAACAGGGGGCCCATACCG	1533
OY	1489	cattctacatcgcgcgcgggagagcggtctaccaagagcgagactctatcttggaaccataa	1548
Db	1534	CTTTCTATCATCTGCCCTCGAAGAAGGTTTACAGGCCCATGCTTTCACCAATAGACTAGTA	1593
OY	1549	accacacatctgcgttatcttaagaagatgycattct---cttaagaacatcataitaaggcca	1605
Db	1594	TCCATTACATGTTGGATTATGAAGATGACGACTACAGACACAGATTAAGCTAAAGTAGAG	1653
OY	1606	aagcagfytgaacatgaacaacacgctagccacgfyggcgagcgctgtgtacgttgaacctga	1665
Db	1654	TAGCAATTGACAGACACAGATNAACATCAATGCGGGAAGCAATTATTAACTTAACCTTA	1713
OY	1666	tcaactgtgaacagacgaactcccactctcttgagagagacgacgtaccgcgcgttttaaag	1725
Db	1714	TCAATTTGGAAATGATGACATCACTATATTTGCACAGAGAGCCCTTACAACTGACATTTGAGG	1773
OY	1726	agacgcgcgagcgccgcgtctccgfyggccaagfygtctgcctctcagacagacatcgaqg	1785
Db	1774	AGACGGTCTGGTGAATGGCTTCCATATTGGTAAATACCGGGCTAAACAGACATCGGTG	1833
OY	1786	acgtagtagtgcattcatctgaayggcaagcgtgtgaaccactgtlcatagatgaatcaa	1845
Db	1834	ACATATGTCGAGCACTCGATATTGGGCAACGCTGCAAACTCCTGAGAAATTGCATAGATA	1893
OY	1846	cggagagatctctcgtfyagatagtaagatfgccttcgactaccacgcgaagaacactcat	1905
Db	1894	CTGAGATGTGTATCGTCTCACGGGAGCGATTACTTGTATTATCAAAGACAGAAACAAATCA	1953
OY	1906	ttgtctaggtgagcgctgcgaagacttcttgggcgagcgcccaacaacaagfyaacacc	1965
Db	1954	TAGTTCAGATTTCTGGCTGTGTGATNACACTAG---TTTACCTTAGAACACAGGCTACCAAC	2010
OY	1966	agctgygatagaaactfygagatgltcaacaacaactcctccacccttaogctgtgccccgt	2025
Db	2011	AGCTACGATATTTTTTGGAGACATATCAACAACGCCACCTATACTGGGAGTGCACAGTT	2070
OY	2026	cgactccaagcgtcagaggaacgcttcccgaaagatacagatalccocggaaaatcaactg	2085
Db	2071	CCAATCCAAAGTGTAGAGAAACAGTTGAAGTCCGGGACCCTGATTTACCGAGGGCTTACGG	2130
OY	2086	ctaccgacccggaaacccaagcgtactcgtggttgaagltgaactgggaactccaccctgg	2145
Db	2131	CGACAGACCCAGACACCAACAGCCGATTTCACCTTCAGATTCATTTGGGCAATTTTACG	2190
OY	2146	ccacacaacgaag---cagaagadacaacccctactyaatagctgggtgtatagtlacg	2202
Db	2191	CTACGAACACAGGACCAATATGAGAACCAACAACACTGCACATACACAGGATCCGTAAATCC	2250
OY	2203	aaacgatataccccaacccgaggaacaccgggttcgcgaatcgggcgcctcgtgtgtcaag	2262
Db	2251	TGACGGTATACCCAGATTCGACAAATACACGGGAGACCTAAGGGTCACTTGGGCGCACGTG	2310
OY	2263	agatccgggaagaagtacatcatcgtactctgaggaattcgaagatgcttaactccacgcgc	2322
Db	2311	AGGTCAGTGAAGGGGTACCATCGATTACAGAGATTTAAGGTGCTGATCCTGTCGCA	2370
OY	2323	gcgtfyaggacctaaacagctcatcggagatgaatacagatbaagcgagcttcaagatca	2382
Db	2371	GGGATATGATCGCAACAACCTGTCTTTGGCCCTATATATACGAACGAATGCTGACGGTGA	2430

[illegible]

Db	3511	TCGTGATTTCATTGCGACGACTACCACTAACCGTCAGGCCATACCAACTTCCTCAATCACGG	3570
Qy	3502	tgctgcgttcccccgaaccacccgacccgctccttcgcgctttctaagagacgcgcgaacaagaag	3561
Db	3571	TGTTTGTTGTGCCAACACTCTTGCTGCCTCAACACTCAGGGCTTTCTTAGGAGCGCTGCTATATGCTA	3630
Qy	3562	gcgcgcgtctcggcgagcgctcgcacac-----gagttccctcgagccgatctacgcga	3612
Db	3631	ATGGTATCTGCTGCTCTGCTTAATATGCGCAGCGAGAGTTCCTCGACAGACTCTCTGGCCA	3690
Qy	3613	ccgaacgaggaagcgcgtccacagcggcgagcgctcgcgttccacgctccagggaaatctggaggg	3672
Db	3691	CTGATGAAGATGGGCTACACGCACGCAGAGTAACCTTCTTCATATGCTGGAAACGATGAAG	3750
Qy	3673	ccggttcgactcttgatatactgaactcgaactcggagcagagaaatcagcggcagctatact	3732
Db	3751	CTGGCGGATATTTTAAATGTG-----TGACACAGCTGACAACTCAGCAATGCTCAGCC	3804
Qy	3733	taacgcagcctttcccccagagcaaatcagacaattcagatcagatccgcgtccgcacagcga	3792
Db	3805	TGAACCAAGCATTCGCCCGCTGCGGCTCCAGCAGTTTGATTTGGTTATGCGGCCACAGGAC	3864
Qy	3793	gcgcgacgagacccgcgcgctttggacgcgaactcgcgtcttcggtggtcttcgtaacca	3852
Db	3865	GGGGACGACGCCGGACCTTAGAGATACCCACATGCTCCCTCACTGCTGGTGTGTGATGA	3924
Qy	3853	cacaaaggagaccgaacttttaagcgaaaatcgaactcactctgcgcttcttcgagggtag	3912
Db	3925	CGCATGGAGACCCCGTGTGTGACGACACACGACTCTCTCTCGCTTCGTTGTAAGAGAG	3984
Qy	3913	aaggcctccgtcgaagagtttgaagctgcgcgaacgaagaagaccttaaaacacacctctgc	3972
Db	3985	CTGGATGTCCGAATAAGTTTCAGGTGCTCAGGCCGATGACCCCAAAATACAGGTGTA	4044
Qy	3973	aagatgaactgcgaagatatctactcaagttatttgaagcgaacaaagaggtctcttcg	4032
Db	4045	TGGAGACTGCGCAATACCATCTACTACTATCTGTTATGGAACATGATGGTACACTTTCG	4104
Qy	4033	tactggaccagtaagaacagctcatcccttcgcgaaagcttggacgcgcgagatgagcgca	4092
Db	4105	CCGTGACGCCGGAACCTTAGCTGATCTATTTCCTGAAGCCCTGGACGCGACCAACACG	4154
Qy	4093	cgcttcaacagctcgcacatcgcgcgcgcgaactcgcgcgcgcacactcgtgacccctctgc	4152
Db	4155	AGCAATACAGGGTGTGGTGGCGCTTCCAAACGCGCTGGGGCAC---CTCCACACTTGT	4221
Qy	4153	agaactccatccctcgttgttaacggttaactatgtaagaagcgaaacccgcgcaccaatttcg	4212
Db	4222	CCCTCTACCTCTCCACCGCTACCATCATGGGGGTTGAGAAACCAACCTTAGACGATCTTCG	4281
Qy	4213	agaagagacctttaaagcggcgagcttgcgaactcgttgaagcagcttgccggaaattgctta	4272
Db	4282	AAAGTAATTTTACACAGCTGGCGCTCTTACACACCGCATATGCAACAGGACGCTGTTT	4341
Qy	4273	ctgcacagcgcgagccacaagaagaagacaacatcacytaacccaatagacgcttcgcgca	4332
Db	4342	ACCTGGGGGCAAAACATTCAAGAGGGCTTCATATGCTGCTACTCGATATATCAAGAAACCA	4401
Qy	4333	tcgcagcttgagcagcagcactagaagccgctgcgcgaactcgcgcttcgcgtcgaatgcgaca	4392
Db	4402	TGAATAATGACAGATGCTGTCACAAAGTTGTGGAGACGCCCTTCGACATTTAACTCTCAA	4461
Qy	4393	ccgcgcgcgtcttcgctcaataatcgaacccaacgcgttcatatgaagggcaatgttgaagtgc	4452
Db	4462	CCGAGATCATATGCTGAATCTTCCACCCAAACTCTGTCTATGACGGCAGTTTGCACTTCG	4521
Qy	4453	acgtcatcgtatacgaatacagcttctgcgaatgacaacagcccggttgaagaatctcactca	4512
Db	4522	AGGTGATGGCTAGTGAACACCGCGTGAAGCAATGATCGAAGCAAAAGTGCATTTATCATGA	4581
Qy	4513	ttctatcgcaaaacgcgltgacccctcaatttctgataaccaaacttgagacggttgaagcaga	4572

Db 4582 TATGACTGCGGTTAGAGTAGCTTCTCTTCTACAAACGGAAGCTGAGTTAAGAGA 4641
 Oy 4573 acagaattatcatagcgcccggttcacaccgggttcaaatgacgtgcaacatcgacc 4652
 Db 4642 GAAAGAAATTTATTCACAAACGTTCCGCAACGGCTTTGTATGACATGTAACATAGACA 4701
 Oy 4633 aggtgtgtccgttcacagacagcgcggtgagcgcaagaagacacacacagagtgccg 4692
 Db 4702 GCGTGTCCGCGGATCGACAGCCGCGATTCGCGAGGAGGATACACAACTCCAGG 4761
 Oy 4693 cgcactcatccggaacaagctgcccgtgacgaacaagagtgagcgctccgacg 4752
 Db 4762 CTCACCTCATACGAGACGACAGCCGCGGTCCAGCCGATATTATGAGGATTAATTACGG 4821
 Oy 4753 acaaggtgtctgtcgacacatccagctgacgtgacgacacacacacacacgctgtgcacg 4812
 Db 4822 AACTCAATATTCATTTGGTGTGATACATGAGAGTACTGATGACATTAATGACGCTACTGG 4881
 Oy 4813 accgtgtacgggtgac 4872
 Db 4882 ACTTTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4941
 Oy 4873 cactagccgacgtctcgcgtgtgtgtaggtctctctctctctctctctctctctctctct 4932
 Db 4942 TCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5001
 Oy 4933 tctgtaggaagaagacactggaacggcgacgtgcaagacacacacacacacacacacac 4992
 Db 5002 TCATATGAGAACGGAACATGGAACCGGCGCATCGAAGCCCTCAACATCAAGATGTTCTTA 5061
 Oy 4993 cgtgtgacgtcggtgac 5052
 Db 5062 CCGACATGAGACCAACCAACGCGGTAGTACAGTCTTAACATTAACAAAGCAGACAGAAC 5121
 Oy 5053 aggtgtccac 5112
 Db 5122 CTGCTTCACATCCCTTTATTAACCGGATGTTAAGACACCTTAACCTAATTAACAGC 5181
 Oy 5113 acgcaagtgagcgac 5172
 Db 5182 A-----AGTATCCCATGACCTGCTTGTGATGCAAGACTTGGAACAGTGTGAAAGATT 5235
 Oy 5173 acttcccgccggcgacg 5189
 Db 5236 ACTTCCACCCGAAAC 5252

RESULT 14
 AF367363 7799 bp mRNA INV 12-AUG-2001
 DESCRIPTION Heliothis virescens truncated cadherin-like protein mRNA, complete cds.
 ACCESSION AF367363
 VERSION AF367363.1 GI:15149241
 KEYWORDS tobacco budworm.
 SOURCE Heliothis virescens
 ORGANISM Heliothis virescens
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 REFERENCE 1 (bases 1 to 7799)
 AUTHORS Gahan, L.J., Gould, F. and Heckel, D.G.
 TITLE Identification of a Gene Associated with Bt Resistance in Heliothis virescens
 JOURNAL Science 293 (5531), 857-860 (2001)
 PUBMED 11486086
 REFERENCE 2 (bases 1 to 7799)
 AUTHORS Gahan, L.J., Gould, F. and Heckel, D.G.
 TITLE Direct Submission
 JOURNAL Submitted (31-MAR-2001) Dept. of Genetics, University of Melbourne, Parkville, Victoria 3052, Australia
 FEATURES
 source 1..7799

CDS
 /organism="Heliothis virescens"
 /db_xref="taxon:7102"
 /note="r1 allele"
 194..2062
 /note="truncated Calp"
 /codon_start=1
 /product="truncated cadherin-like protein"
 /protein_id="AK85199.1"
 /db_xref="GI:15149242"
 /translation="MAVDVRIITAAVLIILAAVLITLPAQDSYVAIIPRPDPDPNPF
 EGVPMSONPLLPEDREDYCMNAFDSALNPVTITMEBIEEDVAIAIANRGNTPT
 TVVTPNFTGTHLGPVIRIPEQGDMLVITRODYEIPNMQTIENVRVEDERFQ
 ATVMILINVIDNAPMIOMEPDIDHEGTTECKSVSDAGSISTRFMTFES
 DRNDEEFELVRENIOGMVYHMRLLINKPLDYEENPLIFRYALDSIPNVHTVM
 MGVENIESRPPEMELFAVOQDEKTAOSFVRAIDGPGIDKPIFYRIETESBKD
 LSEVENTIGRREGAMFKVAPIDBDTLEKVFHSLIAYRGDNVDSGSPSKNDIV
 IIVNDVNDQAPVFPFRPSYFIEIMEETAMTLNDEBFHBDLQPHQIVHLESHPA
 GAHEAFYIAPEVGYQROSFVQNHMHMLDFEVPEFQKIQLRVAIDMDPVRVGIAT
 ININLINMDELPIFEHDVQTVTFKEEGEFVATVLAQRNIDIRVESHSLMGALN
 YLSIDKDTGDLVITIDAFYHQNELFVQIRADDTLGEPTNTACSLSHRGL"
 1982..2007
 /note="r1"
 complement(2029..2052)
 primer_bind
 /note="R3"
 2030..2281
 /note="LtrR"
 2030..4292
 /organism="Heliothis virescens"
 /db_xref="taxon:7102"
 /transposon="Hel-1 retrotransposon"
 /note="r1 allele"
 2221
 misc_feature
 /note="polyA site for 2.1 kb transcript"
 4036..4292
 /note="LtrB"
 4235
 misc_feature
 /note="polyA site for 4.4 kb transcript"
 complement(4322..4351)
 primer_bind
 /note="R2"
 2100 c 1921 g 1791 t
 BASE COUNT 1987 a 2100 c 1921 g 1791 t
 ORIGIN
 Query Match 24.8%; Score 1363.12; DB 3; Length 7799;
 Best Local Similarity 64.4%; Pred. No. 2.8e-230;
 Matches 2170; Conservative 0; Mismatches 1158; Indels 42; Gaps 7;

Oy 1952 cacaagtac 2011
 Db 4288 CTCGAACACCTGCCCAACGCTGATACACTGCAAGACATCAACACACCTCTACCGCT 4347
 Oy 2012 acgctgtcccggttcac 2071
 Db 4348 CAGACTGCCCGCACACGCCGCTCAGTGAAGAGAACGTGGACGGGTTCTGTATCC 4407
 Oy 2072 ccggaatacactgtctac 2131
 Db 4408 CACCGAGCTGCACGCGCTCCGACCCCGACACACCGCGCTTACACATCGATG 4467
 Oy 2132 ggaactcac 2191
 Db 4468 GCACACTTCTCCGCGCACCAAGCAGGAGATCTATGTAAGATTGTAATTG 4527
 Oy 2192 tatagtctgaaagatatataccacacacacacacacacacacacacacacacacacac 2251
 Db 4528 TATAGAAATTCGAGCGGTATACCGAATTTGACACACGAGGACCGCATCGCGCGCT 4587
 Oy 2252 cgtgtgcaagagatccgggaac 2311
 Db 4588 GGTGGTTGCGAGATCCGGGAACAGCTCATATGACATACAGAGATTTCGGGCTCTGTA 4647
 Oy 2312 cctcacgctcgcggtgagacactcaacacacacacacacacacacacacacacacacacacac 2371

Db 4648 CCTCAGTCAGGGTCACGGATCTCAACAGGTCATTGGACGACTATGATATATCAAC 4707
 QY 2372 gtccagatcaacaataatgacatgaacgaacgagccgactctcgaaagcgacgct 2431
 Db 4708 ATTCAAT 4767
 QY 2432 gacgcagacatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2491
 Db 4768 GACGCGAGGATTCGGGTGGAGAGTCCGCCCTCAGAGGATATATATATATATATATAT 4827
 QY 2492 cgcacacgacatgacgagccgctctacaacaagtgcgtacataatacaactaga 2551
 Db 4828 CGGTACCGATATGACGGAGCTCTTATATATCAAGTGCCTATACATCACTCTGAGCT 4887
 QY 2552 caaacctccgaggaatgagtgaagaatgactcacaactgctcacaactgagtgatgc 2611
 Db 4888 GGACACTCCAGAGACCTTGTGAGATCACTTCAATTTGGGTCATATCTCATGTAGAA 4947
 QY 2612 gaacgagcgatgacatgacgagcaacccctgctctactctactctactctactctac 2671
 Db 4948 GCACGAGGCTATGACGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5007
 QY 2672 tagcagcgagctgctccctggaacacgcgacggaatgctcctcagatccaactctga 2731
 Db 5008 CACGCGCAAGTGGACCTCTCTCTGTCACGCTGTCGCGCTGATCTCAACTACTTCA 5067
 QY 2732 agtcacgacgatatgaaatagaatcatcgcacacacacacacacacacacacacacac 2791
 Db 5068 CACGCGCGGAGATATACATCCACATTAACAGACAGCAATAGGTGCTTCGAGTGA 5127
 QY 2792 cactgagaaatcaacacgctgctgctgctgctgctgctgctgctgctgctgctgct 2851
 Db 5128 AGAGGACAGATTGAGGAAAGTGTATATATACGAGGCGCGGAGAGAGAGAGAGAGAG 5187
 QY 2852 gctcagcgtgactcccaacgacgcggaacgagagagagagagagagagagagagag 2911
 Db 5188 CGGCACTCTTCTGCGAGGATCTGATAGAGTGAATCTACCAAGTGGTGCATCA 5247
 QY 2912 gatacacttgcgtaaaccccgactgcggaattcttcgagtgagctggaactgcg 2971
 Db 5248 GATCAACTACGGGATACCTCTGTCGCGCTTCTGAGAGTGAAGTGGAGAGAGAGAG 5307
 QY 2972 tgccttgagtgatcattccgggggagcaaaaattgacacgcgactgagtgacacac 3031
 Db 5308 CCGGTGATACGATCAACACACTGCCGGGAGAACTCGACAGAGAGAGAGAGAGAGAG 5367
 QY 3032 acaactactctgttaaatcactcactctctctgctgctgctgctgctgctgctgctgct 3091
 Db 5368 GATCGATCTCTTCAACGATCATCACTATCACTATGAGAAAGAGATGAGCAACCGGA 5427
 QY 3092 cagagacgaagtgaataatttgctgctgctgctgctgctgctgctgctgctgctgctgct 3151
 Db 5428 TCAAGGAGAGACACAGATGT 5484
 QY 3152 gcaattgctgtaacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3211
 Db 5485 ---ACTGCGTGAAGGCTCTCTCAATGGGATATCTGAGGGCTTGCTACAGGGTCCGTGT 5541
 QY 3212 actccagaatctacgacgagcaaggaatgaacagacagacagacagacagacagacag 3271
 Db 5542 AACCCGATATCTTCCGCCGACCGACGACGACGACGACGACGACGACGACGACGACGAC 5601
 QY 3272 ttaaggaatcccgagctacgacacgagacagacagacagacagacagacagacagac 3331
 Db 5602 GTACGCAATGTCAGCTCAGCGCCACGACAGAGGACATCACTTCCCAACTCTTCAAC 5661
 QY 3332 catgactctgatatga-----acaaactggggaactggaacgact 3376
 Db 5662 CATGATACCATAG 5721
 QY 3377 ggaacttgagggtatctgggacactaagaataltcatctgagccttcgacacagagcta 3436
 Db 5722 GATTTAAAGGCTATTGGGACATTATGAGTACATGCAAGGATACGACCATGAGT 5781

QY 3437 cccgacgagagctcaacgagacgtacacccctggctcactccgcccctacacactcaaca 3496
 Db 5782 TCTCTAAGAGATATCTTACGAGAGATACCCGCTAGTTATTAACCTTACACCTCCACGA 5841
 QY 3497 cccctgtctgctgctcccgcaacccgactccgctcactcgcgtcttcctagggagcgcaac 3556
 Db 5842 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5901
 QY 3557 agaaagcgagctctgagcgagctgacgaagatctcctgagccgactcaagcacaaga 3616
 Db 5902 AGTGAAGCGGCTGCGGACAGAGTGGACGCGGAGTTCCTCGAGCAATGTCGCCACGA 5961
 QY 3617 cgaagagcgccctcagcgagcgagctcagctccacgctccaggggaatggagagcgct 3676
 Db 5962 CAGAGATGCTTACACGCTGAGATTTACCTTCTATCTGTGAGATGAGAGAGGCTT 6021
 QY 3677 tcaactcttgatataactgaagtgagcaagagaanaataagcggcagcttatatacg 3736
 Db 6022 GCAGTACTTCGACGCTGTTAAACGAGGAGTG-----AACTTAGGCGCTGACCATCAC 6075
 QY 3737 cgaactctccagagcaaatagacaatcagatcaagatccgggacagagcgagcg 3796
 Db 6076 GCAGCTCTTCCCTGAAGACTTCCGAGATTTCAAGTGAAGATTCGCTACGAGATGTGG 6135
 QY 3797 cagagagcccgcccgcttgacacgctcaagcttctcggtgctcctgtaaccaaca 3856
 Db 6136 TACGAGACCGAGTCCAAAGATACGAGACTGACGCTGACCTGATGTTGTTCTTACGA 6195
 QY 3857 gggcgaaccagtgctcagcaaaatgcagctactgctgctcctctcgaagtgaaag 3916
 Db 6196 GCGAGAGCTGCTGTCGAGACTACGACCTACACGCTGCTGCTTATTTAGGAAGAGCTGG 6255
 QY 3917 cctcgtgagagtttgagctgcgcaagcagaagacttaaacacactctggaaga 3976
 Db 6255 TATGGAAGAGCGGCGCACGCTCTCTGCGCAAGACCGCGCTTACATATGTGAAAG 6315
 QY 3977 tgactgcgaatctatctactaagttatgacgcaacacagagagcttctgact 4036
 Db 6316 TATATGTCACGACTTATTTACAGCATTTGTGAGGCAACTGATGAGGCTACTTGCGAGT 6375
 QY 4037 ggaacgactcaacagctcactcctctgagagagtgagacgagagtgagcaagctc 4096
 Db 6376 AGACCCCGAGTCCAAAGAGCTGCTCTGCTGACGCGGCTGAGGCGCGGAGACAGAGAG 6435
 QY 4097 ttaacgctgcaatcgcgagcaactcgcgcgaacgcgaactgagatccctctgcagac 4156
 Db 6436 GCACACCTCATATGCGGCGCCAGGACTGCGCCAGCCGCGCG---CGTGTCTGACAGC 6492
 QY 4157 ttcacatcgtgtgcaagtgatgaagaaagcaacccgacgacgaacttctgagca 4216
 Db 6493 TTCCACCTCTACTGTTATGTAATGTTGAGAGACCAACCGCGGCAAGTTCACAGAG 6552
 QY 4217 ggaacttcaacagcgagcattcgaacttgagacagcaattggcgggaattgctactgt 4276
 Db 6553 CGCTGTGATACAGCGCGGATCTCCACCTCGACACCAATCAACAGAGCTGTGACTGCT 6612
 QY 4277 cagggcgagccaacagaagagacacatcaagtaacacatagaacgctgagacatga 4336
 Db 6613 ACACCGACCCATTTCAGAAAGGCTCCCGTGACCTACACGCTGATACAAAGATTCATGA 6672
 QY 4337 gctgagcagcagctgaagagcgctgagcagctcgttcgagctgagctgagcagcag 4396
 Db 6673 AGCTACTTCCACTGCAAGGCTGTGAGAGACACCTTCAACTGAAACCTTCAGACTGG 6732
 QY 4397 cgtgcttccgctcaatagcagccacgcttccatgacagcgacatgctcagatgagcgt 4456
 Db 6733 AGTCTGACCCCTCAACTTCCAGCCACAGCATATGACAGGCAATGTTGATGATGATG 6792
 QY 4457 catgctacgagatcagcttgcgaatgcgaacagccgctgagaaagtctacatcctc 4516
 Db 6793 GATGCTACTGATACAGTGGAGAGACCGCGGCACTGAAGTGAAGGCTGATCATGATTC 6852

```

OY 4517 atcgacaacgcgctgacattcatttcgataaccacttgaccgttgacgacaagc 4576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6833 CGACCGCAACAGAGTGTCTTACCTTCATGACACGCTCGAAGAGTCGAAACCAAGA 6912
OY 4577 aaatttcagcgcacacgttcagcaccggttcacacatgacgttcaacatcgacagc 4636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6913 GGAATTTCATAGGGGAACATTATACCTCTGTTTTCGCATGCGGTGCAACATGACACAGAC 6972
OY 4637 ggtgcgcgttcagcga---cagcagcggtgtgtgcgacaagcaccacccgaggtgcgc 4693
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6973 GCTGCCCCCAACGACCCCGCCGCGCCGCGGAGGACGACCAACCGAAGTCAGGGC 7032
OY 4694 gcaattcatccgcgacgaacgctccgttcagcacaagaagtcgcagccgtccgcagcga 4753
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7033 ACACCTTATACGCGACGACACCTCCCTGTACCGGCTGAGGAGATTGAACGTTTCGGGGTAA 7092
OY 4754 cagcgtgtcgttcgcaccatcagctatgtctgagcaccacaacagcgtgtgtcgaaga 4813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7093 TCCCAACACTAGTGGCGGACATCCAGAACGCCCTCGAGGAGGAGAACCTGAACCTAGCCGA 7152
OY 4814 cctgtgtgcggttgacacatccgacgttagcgagagatcaatgacgacatcgccgtctacgc 4873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7153 CCTGTTCAAGGGGCGAGACTCCATCTTGCGGCGGAGGCGGCGGCGGTGTACGC 7212
OY 4874 actaacgcgcgtctcgcgtgtgtctagcttccctcgtcgtctgactgtcttcgcattgt 4933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7213 GCTGGCGGGGGTGGGGCGCGCGCTGCTGTGTGTGCTGATGCTTATCTTCTTCTT 7272
OY 4934 ctgttgagacaagacgacttgacgcgcgcgtcgaacacatcctcatgacgaagtaacgctc 4993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7273 CATCAGGACTAGGGCCCTCAACCGTCTGTGGAAAGCTCTCTCATGACCAATGACAGTTTC 7332
OY 4994 ggtgacacgcgcgttgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7333 CCMAACACTGGGTGTGAACCGGTGTGGGGGCGGCGGCGGCGGCGGCGGCGGCT 7392
OY 5051 cgaagcgtcccaaccccatgtgtgaacgagccatccgcgcgcgcgcgcgcgcgcgcgcgc 5110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7393 GGAGGGGCTCAACCAATCTGGAAGCAAAACCTCAAGGACCGGACTTTGATGCTTTAG 7452
OY 5111 tgaagcgagttgagcactcgacactgacgtacgacgacgacgacgacgacgacgacgacg 5170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7453 CGAGAGTCGTAACGACTGGATCTGATCGGATCGGATCGGATCGGATCGGATCGGATCGG 7512
OY 5171 ctacttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7513 CTACTTCCCGCTGACGAGGAGAGCTCCATGCGGGAGTCTG-----CAATGAACACAT 7566
OY 5231 cagcgacaacaagccgctgacacgacgacgacgacgacgacgacgacgacgacgacgacg 5290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7567 GCCTGAGACTAATTCAGTAGCAAAACCAATTAATTCGGGTTCAACGCTACCCCTT 7626
OY 5291 cctgcacagc 5300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7627 CAGCCACAG 7636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
LOCUS AE004760 11419 bp DNA BCT 30-AUG-2000
DEFINITION Pseudomonas aeruginosa PA01, section 321 of 529 of the complete
genome.
ACCESSION AE004760 AE004091
VERSION AE004760.1 GI:9949520
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa.
Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
1 (bases 1 to 11419)
REFERENCE
AUTHORS Stover,C.K., Pham,X.-Q., Erwin,A.L., Mizoguchi,S.D., Warrenner,P.,
Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M.,
Garber,R.L., Goltry,L., Tolentino,E., Westbrock-Wadman,S.,
Yuan,Y.,

```

```

TITLE
JOURNAL MEDLINE
REFERENCE
AUTHORS Stover,C.K., Pham,X.-Q., Erwin,A.L., Mizoguchi,S.D., Warrenner,P.,
Hickey,M.J., Brinkman,F.S.L., Hufnagle,W.O., Kowalik,D.J.,
Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E.,
Westbrock-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N.,
Folger,K.R., Kas,A., Larbig,K., Lim,R.M., Smith,K.A., Spencer,D.H.,
Wong,G.K.-S., Wu,Z., Paulsen,I.T., Reizer,J., Sater,M.H.,
Hancock,R.E.W., Lory,S. and Olson,M.V.
Direct Submission
JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics,
University of Washington Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
FEATURES
source
1..11419
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
211..537
/gene="PA3385"
211..537
/gene="PA3385"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC06773.1"
/db_xref="GI:9949521"
/transl_table="MRPLKATPTYSRTADKFEVRLPEGMRQIAPVARSRRMS
ETIRLEOSLQEBALDNLGVRDSELSHERELLORROLTHROQNLVALLIAD
AELAQ"
691..1056
/gene="PA3386"
691..1056
/gene="PA3386"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC06774.1"
/db_xref="GI:9949522"
/transl_table="MSYRLADALVWHLGFLFVFGLLPLRPLWHLPAVAV
GCAVEFEGPEPLPLENRLRRAAGDAGSGFEVHLPLIYAGLTPAVQWILAI
VLIVNLAYLYLMHCRSR"
1154..1924
/gene="rhlG"
1154..1924
/gene="rhlG"
/codon_start=1
/transl_table=11
/product="beta-ketoacyl reductase"
/protein_id="AAC06775.1"
/db_xref="GI:9949523"
/transl_table="MHPFSLAGRIALVTGSSRGIGOMIAGLLEAGARFICARDAE
ACADTATRLSYGCGQAIIPADLSEBARRLAQALGELSRLDILVNNATSGKALE
SYFVSGWEKQWLNTSVSCTQLDPLRLRSASAEHPARVINTGSAVAGISAMEQAY
AYGSKALHOLSRMLAKELGHEINNVLAIPGPPSRMRHIANDOALEASASIP
MGRWGPREEMALAIISLAGTGAAYMTGNVIPIDGGFHL"
1977..2672
/gene="PA3388"
1977..2672
/gene="PA3388"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC06776.1"
/db_xref="GI:9949524"
/transl_table="WTHSVSPITGYIRSCFEKFAIPROPPLAPARGTLELPPEDV
EALBGLHGVSHVWLLFLHQALBDEKPLKVRPRLGNGRSLGVFATPATRHPNGIGS

```


[illegible]

Search completed: January 15, 2002, 16:00:37
Job time: 8795 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:34:47 ; Search time 4372.01 seconds
(without alignments)
13513.302 Million cell updates/sec

Title: US-09-715-909-1

Perfect score: 5498
Sequence: cataatacataataagaagga.....aaaaaaaaaaaactcgag 5498

Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estpl:*
5: em_estpl:*
6: em_estlba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	62.8	1.1	1100	13	CNS016KD
C 2	55.2	1.0	925	13	CNS0091P
C 3	53.4	1.0	513	10	AA948367
C 4	53.2	1.0	959	11	BC387114
C 5	52.8	1.0	300	11	N98097
C 6	52.8	1.0	656	10	BE602179
C 7	52.2	0.9	314	10	AU198049
C 8	51.8	0.9	1043	13	CNS0145P
C 9	51.6	0.9	523	10	AV401734
C 10	51.4	0.9	519	11	BI140996
C 11	51.4	0.9	704	13	CNS062FV
C 12	51.2	0.9	945	13	CNS01107

C 13	51	0.9	503	10	AL513809
C 14	50.8	0.9	482	11	BF765609
C 15	50.8	0.9	524	11	BF762445
C 16	50.8	0.9	726	11	BF819608
C 17	50.8	0.9	736	11	BG747190
C 18	50.8	0.9	807	10	AU124569
C 19	50.8	0.9	1101	13	CNS002M0
C 20	50.6	0.9	512	10	AV401445
C 21	50.6	0.9	556	10	BE510534
C 22	50.6	0.9	1068	13	CNS04BN6
C 23	50.4	0.9	439	10	AV637890
C 24	50.4	0.9	1125	10	AL547503
C 25	50.2	0.9	1200	11	BG438325
C 26	50.2	0.9	245	8	BF048950
C 27	50.2	0.9	782	11	BF631315
C 28	50	0.9	195	11	BG156792
C 29	50	0.9	599	13	AQ160671
C 30	50	0.9	910	13	CNS01T91
C 31	50	0.9	935	13	CNS006XK
C 32	50	0.9	1201	13	CNS00Z07
C 33	49.8	0.9	210	10	AU060678
C 34	49.8	0.9	245	10	AU056151
C 35	49.8	0.9	604	13	AA550384
C 36	49.8	0.9	936	13	CNS003RX
C 37	49.8	0.9	957	13	CNS015W7
C 38	49.8	0.9	1184	13	CNS04PAP
C 39	49.6	0.9	510	10	AV405247
C 40	49.6	0.9	512	10	AV402969
C 41	49.6	0.9	700	10	AL507642
C 42	49.4	0.9	500	10	AU087313
C 43	49.4	0.9	927	13	AZ046250
C 44	49.4	0.9	1101	13	CNS00E7Y
C 45	49.4	0.9	1101	13	CNS0180F

ALIGNMENTS

RESULT 1
CNS016KD/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit
fly) genomic survey sequence.

ACCESSION
AL106855
VERSION
AL106855.1 GI:5624152
KEYWORDS
GSS.
SOURCE
ORGANISM
Plasmod Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 1100)
Genoscope.

AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EGGP) -
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
location/Qualifiers
1..1100
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN16D22"

QY	1621	acaacgctagccacacgctgagcgagcgctggtgtatagctgaaactgtatcaactgtgaacagca	1680
Db	1621	acaacgctagccacacgctgagcgagcgctggtgtatagctgaaactgtatcaactgtgaacagca	1680
QY	1681	aactccacattctgaaagaaagaaagcttaccgcgctgtcttaagaaagacgttcgcgcgcg	1740
Db	1681	aactccacattctgaaagaaagaaagcttaccgcgctgtcttaagaaagacgttcgcgcgcg	1740
QY	1741	gcttcoccgctgagccacgctgctcgcctctcgacagaaacaatcgacagcctagtatgagcat	1800
Db	1741	gcttcoccgctgagccacgctgctcgcctctcgacagaaacaatcgacagcctagtatgagcat	1800
QY	1801	cattgtatgagcgaaacgctgtgttgaactaccgttcaatagaatcaacgagagagatcttcg	1860
Db	1801	cattgtatgagcgaaacgctgtgttgaactaccgttcaatagaatcaacgagagagatcttcg	1860
QY	1861	tgaagcatgagaaagcttcctctgactatcacacgagaaacactatatttcttaagctgcgcg	1920
Db	1861	tgaagcatgagaaagcttcctctgactatcacacgagaaacactatatttcttaagctgcgcg	1920
QY	1921	ctgacagatattcttggcgagacgagcccaacaacaacagctgacacacccagcttggctgaatagac	1980
Db	1921	ctgacagatattcttggcgagacgagcccaacaacaacagctgacacacccagcttggctgaatagac	1980
QY	1981	tggagagatgttcaaaacaacatctctcccaacccctagctgtccgcgttgcgactccaaagctgcg	2040
Db	1981	tggagagatgttcaaaacaacatctctcccaacccctagctgtccgcgttgcgactccaaagctgcg	2040
QY	2041	agagaaagctgtcccgaaagagctagagaaatctcccggaagaaatctctacacgacccgagca	2100
Db	2041	agagaaagctgtcccgaaagagctagagaaatctcccggaagaaatctctacacgacccgagca	2100
QY	2101	ccagagccctaaccttggcttgcgagatctgagacttgcgacccctgcgcgcgcacaaagcagcgca	2160
Db	2101	ccagagccctaaccttggcttgcgagatctgagacttgcgacccctgcgcgcgcacaaagcagcgca	2160
QY	2161	gagagagaccaaactctgtaataatgctgggtgtgtatagttatcgaaacgatatccccacgcg	2220
Db	2161	gagagagaccaaactctgtaataatgctgggtgtgtatagttatcgaaacgatatccccacgcg	2220
QY	2221	agggcaaacacggggtctccgcgacatcgcgcgctcgtgtgtgaaagaaagatccgggaaacaactca	2280
Db	2221	agggcaaacacggggtctccgcgacatcgcgcgctcgtgtgtgaaagaaagatccgggaaacaactca	2280
QY	2281	ccatcgaccttcgaggaattcggagatgcttaccctcaacgcgtccgcgctgagggagactcaaca	2340
Db	2281	ccatcgaccttcgaggaattcggagatgcttaccctcaacgcgtccgcgctgagggagactcaaca	2340
QY	2341	ctgtgcatactggagagatgactatagatgagagcgagcttcaacgatacaataatctgaacttgaacg	2400
Db	2341	ctgtgcatactggagagatgactatagatgagagcgagcttcaacgatacaataatctgaacttgaacg	2400
QY	2401	acaaacgcgcgcgactcttcgcgaaacgagacacgctgaacgacagatgagctgcgcgcgcgctg	2460
Db	2401	acaaacgcgcgcgactcttcgcgaaacgagacacgctgaacgacagatgagctgcgcgcgcgctg	2460
QY	2461	cggccagcgagcaagcttcatctgcgctccgcgtgcgtgcgcacacgacatcgacgcgcgcgtctaca	2520
Db	2461	cggccagcgagcaagcttcatctgcgctccgcgtgcgtgcgcacacgacatcgacgcgcgcgtctaca	2520
QY	2521	accaaagctgcgtacatcatatatacaaaccttaagaaacaccccgagagatagtgtaagattg	2580
Db	2521	accaaagctgcgtacatcatatatacaaaccttaagaaacaccccgagagatagtgtaagattg	2580
QY	2581	acttcacaacatctgttcaaatctgaaagtgtgaatgagcgaaacgagcgatcgatgcagaaacacct	2640
Db	2581	acttcacaacatctgttcaaatctgaaagtgtgaatgagcgaaacgagcgatcgatgcagaaacacct	2640
QY	2641	ggcgctctctactctgtatcatatacaccgttcatcgcttagcgacgagctgctcccttgaaacacgca	2700
Db	2641	ggcgctctctactctgtatcatatacaccgttcatcgcttagcgacgagctgctcccttgaaacacgca	2700

QY	2701	cggaaatgctccctccgaattccaactacttggaaagtctccagcgcatatcgaataagaatacca	2760
Db	2701	cggaaatgctccctccgaattcccaactacttggaaagtctccagcgcatatcgaataagaatacca	2760
QY	2761	tcgaacaaanaaanaagtgctcgagcgctctcaacttgagaagtccaacagaagtggtacg	2820
Db	2761	tcgaacaaanaaanaagtgctcgagcgctctcaacttgagaagtccaacagaagtggtacg	2820
QY	2821	tcctggagaaatgcgacagagcgcgacagaggtggatccagactgtactcccaagacgtgtgaca	2880
Db	2821	tcctggagaaatgcgacagagcgcgacagaggtggatccagactgtactcccaagacgtgtgaca	2880
QY	2881	gagacagatgtgtaccacacggtatcagataacgataaatttgcgtgtgaaccccgcatcgc	2940
Db	2881	gagacagatgtgtaccacacggtatcagataacgataaatttgcgtgtgaaccccgcatcgc	2940
QY	2941	ggagattctctcgaggtgagaccttggaactcgtgcgccttgtagtgattatcccgggggagcg	3000
Db	2941	ggagattctctcgaggtgagaccttggaactcgtgcgccttgtagtgattatcccgggggagcg	3000
QY	3001	aaaaatgtgaacgagatgtagggatgagactcaacaatactactctcttgtaaatltcatgata	3060
Db	3001	aaaaatgtgaacgagatgtagggatgagactcaacaatactactctcttgtaaatltcatgata	3060
QY	3061	aactctcttctcgtatgltgacggttagagaaaccaagagcgaagtgtgaaatattgtcgttc	3120
Db	3061	aactctcttctcgtatgltgacggttagagaaaccaagagcgaagtgtgaaatattgtcgttc	3120
QY	3121	tattggatgtggaagagaacaagctccctcgagatgcatctgcatctgaattgaattccggttttag	3180
Db	3121	tattggatgtggaagagaacaagctccctcgagatgcatctgcatctgaattgaattccggttttag	3180
QY	3181	ttctccgaagagagcagatgtgcgtgtcgtgtactccccaataatctacagcacgcggacaag	3240
Db	3181	ttctccgaagagagcagatgtgcgtgtcgtgtactccccaataatctacagcacgcggacaag	3240
QY	3241	atgaaacagagacaacgacaactcgcgtgtcgtgtacggaatccctgagcctcagatcacgcg	3300
Db	3241	atgaaacagagacaacgacaactcgcgtgtcgtgtacggaatccctgagcctcagatcacgcg	3300
QY	3301	accgagagcatctcgaggtgagccggaatctcttccacatgatactcgatttgaanaaanaaactgggg	3360
Db	3301	accgagagcatctcgaggtgagccggaatctcttccacatgatactcgatttgaanaaanaaactgggg	3360
QY	3361	aactctgagacgcgtactatgaaactctgaggggttatgtgggcactctacgaaatctcatctgag	3420
Db	3361	aactctgagacgcgtactatgaaactctgaggggttatgtgggcactctacgaaatctcatctgag	3420
QY	3421	cctctcgacacagcgtactaccgcgagcagaaaggttccaacgagagcgtataccctgtgtacgcgc	3480
Db	3421	cctctcgacacagcgtactaccgcgagcagaaaggttccaacgagagcgtataccctgtgtacgcgc	3480
QY	3481	ccttaaaacttccacaacacccgtgtgttcgtgttcccgcaacccggaactccgttcatctggcgtt	3540
Db	3481	ccttaaaacttccacaacacccgtgtgttcgtgttcccgcaacccggaactccgttcatctggcgtt	3540
QY	3541	ctaaagagcgcgcaacaagaaagcgcgctctctgtgcgacggtctccaaacaggtcttccgtgagc	3600
Db	3541	ctaaagagcgcgcaacaagaaagcgcgctctctgtgcgacggtctccaaacaggtcttccgtgagc	3600
QY	3601	cgaatctacgcacacagcagagagcgcctccacgcgagcgatcgatcttccagctccag	3660
Db	3601	cgaatctacgcacacagcagagagcgcctccacgcgagcgatcgatcttccagctccag	3660
QY	3661	gaaatgagggggcggtctcagttacttgtatataactcgaagtggtggagcaggaagaaataagcg	3720
Db	3661	gaaatgagggggcggtctcagttacttgtatataactcgaagtggtggagcaggaagaaataagcg	3720
QY	3721	ggacagcttatattacgcgcagcgttctcccaagaaatacagaacaattcagatcaacgattcc	3780
Db	3721	ggacagcttatattacgcgcagcgttctcccaagaaatacagaacaattcagatcaacgattcc	3780
QY	3781	gggcacacagacggtgcgacagagagcccggcgcgcttcttgacacgagctcacggttcttcggtg	3840

[illegible]

Db	1561	cggtacacagcgagacattccattcgtgycacccctcaatccctcaatcgctggtatga	1620
Qy	1571	agatgtaattttccagaacatcatcattaaaggtcaagcgatgacatgaaacagctag	1630
Db	1621	agtgccagagattccagatattacgattccggtgtgtgaagccagcaacagacagag	1680
Qy	1631	ccacgtgagcgagcgctgtgtatcgttgaactgtatcaatctgaaacagcgaaattccat	1690
Db	1681	gcaacgtgtggtctgcgttggttcaacttctacattcaattgaaacagatagacgcgat	1740
Qy	1691	cttcgaagaaagacagattaccgcgttcgttttaagaaagacgcgtgcgcgccttccgt	1750
Db	1741	cttcgaacacgcgcgtgcagacgcgttcacctctgcagcagatctgaagcgaggttctcgt	1800
Qy	1751	ggccacggtgtccgcgccttcgaacagaaatcgacgacgtatgattgattcaattga	1810
Db	1801	cgcaaaagcggtgttcacacacgacagaaacatctcggtggtgtcgtccagacataatttgg	1860
Qy	1811	caacgctgttgtaactctgttcataagataaataaggaaggaagcttcctgtgaagata	1870
Db	1861	taacgcctgtaacttctctgacatctgcacaaacttcacgcgtgcagcatctcgcgtctca	1920
Qy	1871	cgatgccttcgtaacacacgcagaaacacatcthattgtttcaggtgtgcgcgtgcagatac	1930
Db	1921	cgactccttcaactacatccatcgaaagaagtgaattattgttcaggtgtgcagcttaca	1980
Qy	1931	tttggtgagcagcgccacacaaacacaaatgacacacgcgtcgtgtgatagaacttggagat	1990
Db	1981	gctgtgtgcga---accttccacacggtgcagcgcacgcgtgtgtatcaacgactaaatgacat	2037
Qy	1991	caaaacaactctctccaccccttaacgtcttgcgcgttcgcacttccaaagcgttcgaagaaacgt	2050
Db	2038	caaaacaacgcaccccaactttagcgtcgtcgtcgaagcaqctcccaagtgtgaagaaacgt	2097
Qy	2051	ttcccgaaagataagagatatacccggtgaataactcgtctacacgcccgcgcgcacacagctc	2110
Db	2098	gctcgatagcgccacgcgttcaatcccaagagatatacgcgcacacgcgcgcgcgcacacacgcga	2157
Qy	2111	cctgtgtctcgagatcgactgagtcgagacccacacttgggtccacaagcagggcagaaagaa	2170
Db	2158	ctctggtctcgagataaactcgtggacaacactcttctgcaccaaagcgaagtcgcgcagctaa	2217
Qy	2171	ccctactgaatacgttcggtgtatgtatttcgaacagatatccccaacgcgaggtgacacgcg	2230
Db	2218	ccccgcagagatttagaagattcgtgtgaaatctgaaacatcttcccccgcagatataaacacgcg	2277
Qy	2231	gggttcgcacatcgggtgcgtctgtgtgaagaagatccgggtgacaaagtctaacatcgatt	2290
Db	2278	gggtgcgtgcacatcgggtgcgtctgtgtgcgcgcgcgaaatcgaacacaaagctgacatagacta	2337
Qy	2291	cgaagaaatcgaagatgcttactccctacgcgttcgcgcgtgtgaaggtgacctcaaacatgttcatcg	2350
Db	2338	cgaagagatttgaagttctctctccctccacagtgaggtgtgcgttgaaacttaacacgcgttcaag	2397
Qy	2351	agatgatacgaatgaaagcgacgttcaacgtatcaaatcaaatcgacatgaaagaaagacgcgc	2410
Db	2398	agaaacacacgaacgaatcgtctcctccacaataacataacatcgataagaaacacgcgcgc	2457
Qy	2411	gactctcgcaaacgcgcgtgtacgcagacagatgctgcgtgtgcgcgcgtctggccagacgcg	2470
Db	2458	ggtgtgtgtgtgaggtgacatctgcagcaagaacttcgcgcgcgcgcgcgcgcgcgcgcgcgcg	2517
Qy	2471	cacgcctatcgcgtctcgtgtgtctgcacacgcgaatctgcacgcgcgcgccttaaacaaagctgc	2530
Db	2518	gctcgtgtgtgtgtcctcgtgtcgc	2577
Qy	2531	ctaacactatacaacttgaacaaacacactccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	2590
Db	2578	ataaacatttccctcgtgtgaagaaacagataagacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	2637
Qy	2591	tgtctcaaatgtgaagtgtgaatgcgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	2650

[illegible]

[illegible]

QY	4799	ccgtgctcgaagacactgtgtgacgggtgaacactccgaagcgaaggagagtaataTca	4858
Db	4844		4902
QY	4859	gatgcgcgtacgaactaaGccggctctccgtgtgtcaggtccctcgtccgtact	4918
Db	4903	-ctgcaaggggtgaagctctgtgctccactgtctgtgtgtcgtgttccatggtcgttgc	4961
QY	4919	gctctccgcatlgttctgttgagcaagagcactgaacccggcagctgtcaagcactccat	4978
Db	4962	actgcttaacttcatcatcagcagactagmnmgctaaccgaagcgttggaaGcccttgat	5021
QY	4979	gacgaagtacgctcgtgtggtacccggctcgtcgaacggcgccgggtc---ggcgccgggac	5035
Db	5022	gacgaagtacggtctcaactgtactctgtgatTgaacccggcgccatcgcgccccggcac	5081
QY	5036	caacaagcagccgctcgaggggtcccaaccatgttgaacggagggccatccgcgcgcga	5095
Db	5082	caacaacaacactgtgtgaaggctcccaaccttacttcaatgaagaataaagacgcaga	5141
QY	5096	cttcgacgcacatcagttacacggatgtagcactccgacatccgtacgcacatcgagacatgc	5155
Db	5142	tttagatgacctatagcaggggttccaaagcactctgactgcgacatcgacatgaagtctgc	5201
QY	5156	gcaattccgcgacgactacttccgcgcggcgacacaga	5194
Db	5202	gcacttggcaacgctctcatgtactctgagtTgaagca	5240

RESULT	3
AAAF30934	standard; cDNA; 5527 BP.
AAAF30934	
AC	AAF30934;
XX	
DT	09-JUL-2001 (first entry)
XX	
DE	Heliothis zea Bt toxin receptor cDNA.
XX	
KW	Corn earworm; CEM: Bacillus thuringiensis; Bt toxin;
KW	receptor; crystal protein; CryIA; biological control; insecticide;
KW	crop protection; ss.
XX	
OS	Heliothis zea.
XX	
FT	Key
FT	Location/Qualifiers
FT	CDS
XX	162..5315
XX	/*tag= a
PN	W0200136639-A2.
XX	
PD	25-MAY-2001.
XX	
PF	17-NOV-2000; 2000MO-US31674.
XX	
PR	18-NOV-1999; 99US-0166285.
PR	21-SEP-2000; 2000US-0234099.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Flannagan RD, Mathis JP, Meyer TE;
XX	
DR	WPI: 2001-329223/34.
XX	
PT	P-PSDB: AAB20459.
XX	
PT	New nucleic acid encoding Bt toxin receptor, useful for screening and
PT	identifying new Bt toxin receptor ligands useful as new insecticidal
XX	toxins -
PS	Claim 1(a); Page 61-68; 85pp; English.
XX	
CC	The present sequence of is that of cDNA encoding the Bacillus
CC	thuringiensis (Bt) toxin receptor (see AAB20459) of the corn

Db	2028	gacatlaaacacaccctctcttaoagcttaagctctgcctccgcgcacatccatcaagttgaaag	2087
Qy	2046	aacgtctccgaagaatacagagatacccgysnaaatactgctacagacccgcagacccgc	2105
Db	2088	aacgycgcgcagcggtttgtgatacccaacgcagtctgacagacagaccccgacactaca	2147
Qy	2106	gctacacgttggtcttcgagatccagatctggtactccacttggccacacaaagcaggtcagag	2155
Db	2148	gctgcagctgcgcttcgcagatccagatccgtacagaaatctgtatgctaccaaagcagagcagat	2207
Qy	2166	accaaaccctactagatacgtccgggtgtatggttatugaaagcatatccccacgagagc	2225
Db	2208	actgactcttaagaggtatatactcgggttgtagaatacgaagacgatataccagataaac	2267
Qy	2226	aaccgggtctccgcacatccggcgccctgtgtgtgcaagaatccggagacaagttcacctc	2285
Db	2268	cagcgagcgcaacgcacatccgcgcgcgtgtgtatgtgcgagaaatccggagacgggtcaccta	2327
Qy	2286	gacttcgagaaatcgaagatgcttaactcaaccgtccgcgtgaagagacctcaacatgtc	2345
Db	2328	gactatgagatgtttggaagtcttaactataccataccgactatgttagagatctcaacggt	2387
Qy	2346	atcgagagatgactacgtctgtagcggaegttaacagatacaataatccagatgaagacac	2405
Db	2388	atctggagaaagacccaigtatataatccaaatccatccagataatagacatgacagacac	2447
Qy	2406	gcgcggatcttcggaacgcgcagcgtgacagcaatgtgcgtgtgcggaagcttgcgcgc	2465
Db	2448	ctctccctgtggtgtagaaggaagcaccctcgagcgaaagttccgtgtgcgagaaagtgtgcagc	2507
Qy	2466	agcggacacgtcatccgctcccggtgcgcgcgcacacgaatacgaagccgcgtctacaaccaa	2525
Db	2508	tcagagatgttataagatactccgtactgtgcacatgcagatccgacgcgcgtgtataatcaa	2567
Qy	2526	gtgcgcctacactatacaacctagaacaacatcccgagaggttagtgaagatgtgacttc	2585
Db	2568	gtgcggtatatacttactccagactatgacacatccagaaagactctgtgaaataagacttc	2627
Qy	2586	acaactgtctcaatctgaggtgtggaatgcgaacgaagcagatcgatgcagagaaacccgtgc	2645
Db	2628	aaccgggtcagatctccgttaaaagttaacaccaggtctatagacgcgagacgcgcgcggt	2687
Qy	2646	tctactgttactaacacgcgtacatcgtctaaagatgtctccctgtgaaacacgcagagaa	2705
Db	2688	cagaaactctactacaacgcgtacagctagcttagtgaacagtgcactctcttaactgtcaatcag	2747
Qy	2706	gtctctccagatcttcaactactctccgaagtccaagcgcgatatcgaataatgaataccagac	2765
Db	2748	gtctccgctctgacccttaacttacttcttgagacacgggagagatataccataccataacgagc	2807
Qy	2766	acaacaacaagaagtccttgcgcgcgcctacgtcagagtttaacagaagcgtgtlaagctcgg	2825
Db	2808	acgaaacaacaaagtgtcctcaagtctgaagacgaacgaattccagatgcagcgtgtuatacctac	2867
Qy	2826	gagaatgcacagcagcgcgcgcgcgcgcgcgtgtgtccagctgttaactcccaagcgtgacagagc	2885
Db	2868	gagcgcgcgagcagatgtgacaacaactgtcgtgcagatctactacgcagagatctgtatagat	2927
Qy	2886	gagttgttacaacacggtgtacgatacagatgtgaactcttggtgtgagaaaccccgacgtcgagat	2945
Db	2928	gaactctacaacaagaagtgtgaactacacgatacctacacatccagatcaactctgcgtctccgagc	2987
Qy	2946	tctctcgaggtgcacctgcgagcatgtgtcgccttgggtgtgcattacccggggaggaanaaa	3005
Db	2988	tctctcgagatgcacctcggaagttccgcgtccctgttgcgtcaacaacacacccgcgcgcgagctg	3047
Qy	3006	tttgaccgcgattgggtgagcttacaatactactcttgttaattttcatctgataacttc	3065
Db	3048	ctggaacaggtgagcgcgcgcgcacacatacgcatctcttcctaagttcatgcatgataacttc	3107
Qy	3066	ttctctgattgtgcaggttaggagaacacgcgcgaagatgtgaataattgtctgtctctatgt	3125
Db	3108	tatggaaagagatattgccaacgcgaatccgaacgcgaacagatgtcttagttagtattgtctt	3167

[illegible]

```

OY 4191 ggaacccgcgcacatttccagcagacacttaccacagcgagcttcagcgttgac 4250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4233 ggaacccgcgcacactgttccaaagacacttaccacagcttcgtctcgtcgat 4292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4251 agcattgcccggatcttctacgttcagcgccacacaggaagacacacacacg 4310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4293 ttcatcgaataaactctgtcactttagtagcacacatcagaagacttgcctacac 4352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4311 tacacacagacgcgttcgacgacgttcgacgacgacgacgacgacgacgacg 4370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4353 tacactcgtacatacaagatccatgaagacacacacacacacacacacacacac 4412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4371 gcttcgcgtcgtacatgcacacacgcgctgtcttcgtcgaatgatgcacacgcttc 4430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4413 gcttcacccacacacacacacacacacacacacacacacacacacacacacac 4472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4431 atgcacgacatgttcgacgttcgacgttcgacgttcgacgttcgacgttcgac 4490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4473 atgcacgacatgttcgacgttcgacgttcgacgttcgacgttcgacgttcgac 4532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4491 gcccgctgaaagcttaccatccatccatccatccatccatccatccatccatcc 4550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4533 acggaagagaaagtggtactcgtatatacagcgcacacgagtggttcttcaagttcaaac 4592
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4551 caacttgacacgcttcgacgacgacgacgacgacgacgacgacgacgacgacgac 4610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4593 ccaactgctgagatccacacacacacacacacacacacacacacacacacacacac 4652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4611 aacatgacgttcgaatccatccatccatccatccatccatccatccatccatcc 4667
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4653 ggcatagacgttcgaatccatccatccatccatccatccatccatccatccatcc 4712
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4668 caaagacgacacacacacacacacacacacacacacacacacacacacacacac 4727
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4713 aagagacgacacacacacacacacacacacacacacacacacacacacacacac 4772
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4728 caaagagtcgaagcgcttcgacgacgacgacgacgacgacgacgacgacgacgac 4787
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4773 gagagacgttcgaatccatccatccatccatccatccatccatccatccatcc 4832
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4788 agcaccacacacacacacacacacacacacacacacacacacacacacacacac 4847
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4833 gagagacgacacacacacacacacacacacacacacacacacacacacacacac 4892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4848 gagtaacgacacacacacacacacacacacacacacacacacacacacacacac 4907
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4893 gagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4952
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4908 tgcctcgtactcgtcttcgtcgtcttcgtcgtcttcgtcgtcttcgtcgtcttcgt 4967
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4953 gttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4968 gcaacttcacagacgaataagcgttcgttcgttcgttcgttcgttcgttcgttcgt 5024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5013 gctcgtcgtacacacacacacacacacacacacacacacacacacacacacacac 5072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5025 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5073 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5085 cgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5133 aagcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5145 gagacacacacacacacacacacacacacacacacacacacacacacacacacac 5204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5193 gaagacttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5205 ggcacgcgc 5213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5253 ggaagtcgc 5261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 4
AAF30975
ID AAF30975 standard; cDNA: 5527 BP.
XX
AC AAF30975;
XX
DT 23-JUL-2001 (first entry)
XX
DE Pink bollworm Bacillus thuringiensis toxin receptor BT-R2 cDNA.
XX
KW Pink bollworm; Bacillus thuringiensis; crystal toxin; CryIAC;
KW receptor; BT-R2; insecticide; pesticide; biological control;
KW cotton; ss.
XX
OS Pectinophora gossypiella.
XX
FH
FH Key Location/Qualifiers
FH CDS 154..5343
FH sig_peptide 154..213
FH sig_peptide 154..213
FH /tag= b
FH /note= "the signal peptide is alternatively
FH mat_peptide 214..5340
FH /tag= c
FH /note= "the mature protein is alternatively
FH encoded by nucleotides 154..222"
FH encoded by nucleotides 223..5340"
XX
PN WO200134807-A2.
XX
PD 17-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US41521.
XX
PR 26-OCT-1999; 99US-0161564.
PR 24-OCT-2000; 2000US-0161564.
XX
PA (BULLA L. A.
XX (CANDU) CANDAS M.
XX
PI Bulla LA, Candas M;
XX
DR WPI: 2001-335932/35.
DR P-PSDB; AAB82331.
XX
PT Novel Bacillus thuringiensis toxin receptor from pink bollworm, for
PT identification and design of toxin receptors, for use in combating
PT emergence of toxin resistance and producing transgenic organisms -
XX
PS Claim 3; Fig 1A-B: 56pp; English.
XX
CC
CC The present sequence is that of cDNA encoding a novel 194 kDa high
CC affinity receptor, termed BT-R2 (see AAB82331), of Bacillus
CC thuringiensis (Bt) crystal toxin CryIAC. The cDNA was obtained
CC from pink bollworm (Pectinophora gossypiella), the most destructive
CC pest of cotton worldwide. Degenerate primers based on conserved
CC regions of Manduca sexta and Bombyx mori BT-R1/75 were used to clone
CC partial fragments of BT-R2 from insect midgut RNA. Based on the
CC sequences of the partial clones, sense and antisense primers were
CC designed and used to clone the 3' and 5' ends of the BT-R2 clone by
CC RACE. The full-length cDNA was then obtained using gene-specific
CC primers from the 5' and 3' untranslated regions. The minimum toxin
CC binding fragment of BT-R2 has been identified. Isolation of BT-R2
CC cDNA permits the analysis of receptors in pink bollworm and other
CC insects that affect crop growth and development, as well as the
CC development of assays for the cytotoxicity and binding affinity of
CC potential pesticides. The clone permits the manipulation of
CC natural and/or introduced homologous receptors and hence the
CC specific destruction of organisms, tissues and/or cells of a target
CC host, including insects resistant to Bt toxins. Transgenic
CC organisms expressing toxin receptors can also be generated. The
CC BT-R2 cDNA sequences will also facilitate understanding of the
CC molecular biology of the toxin receptors in the pink bollworm and

```


[illegible]

Db	3521	acaattacaatctctctcaagtttcgttttctcgtgaattccgggacgatctcaagcttgctt	3580
QY	3544	gggagcgcgacaagaagcgcgcttctcggacggtcgccaagagttctctgagccga	3603
Db	3581	tggagacgcgacgtgtgaataatgatttttctcactgttaacggtgtacccgtttaagacaga	3640
QY	3604	tcttagccaccgacggcgagcggtctccacggcgcgagcgtcaacgttccaagtcacggaa	3663
Db	3641	tacaagaattctgcgcgcgacgagtccttgagtcgctggcgctgggtgactcttgatctgttggag	3700
QY	3664	atgagagcgccgtctcagtaactctgtatataactggaagctggggacgaggaataatgcgggc	3723
Db	3701	atgctgtgtggtctaaactactctcaaggttaaatatgtagtgcgacagcttctgaaactgtc	3760
QY	3724	agcttatatacgccagctcttttccagagcaaatcaagacattcaagatccagatccagtcctggg	3783
Db	3761	tgtcgtgaaa-----cagcgctctctctgaaagagcaagaaatttgaggttaccatccgg	3814
QY	3784	ccaagacggcgcgacaggaagccggccggcctcttgagccgaagctacagcttctcgttggtct	3843
Db	3815	ctaaagacggcggaacacagaccctcgatcatatcaacagactccactataacagtcctct	3874
QY	3844	tcgtaccaccaacagggcgacaccagtgcttaagcgaaatgycagctactgtctgcttcttcg	3903
Db	3875	tcgttcgcgactttgggtgatactccttccaaagataacacttaccagtaacatcttctg	3934
QY	3904	agggttgaaagagcctcccgtagaagtttggagcttgcgcggaagcagaagaccttaaaac	3963
Db	3935	aaaaagaggtgtgtctgtacgtgaagagttctcgtcccaatgcaagagaccttaagaaca	3964
QY	3964	acccttcggaagatgagactctgccaagatactactacaggttatttgcgcgcaacaagag	4023
Db	3995	aacctgcagcgaagactgtgcagatattacttaacaggaacttttgggtgtgtgattacg	4054
QY	4024	gtcttctcgtactcgtgaccagtgcaagaagctatctccttcgtcgcaagagtttgacccgc	4083
Db	4055	agccatttgcaccttggaccgcgggtgacgaagctgtgactctcttgaaatcgaabactagaccgg	4114
QY	4084	agggtgccaagcttcttacaagcttgcaatcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgc	4143
Db	4115	agacacactgctacgcagtggtgtgcaagtgtagcagcgaatattctgcaccaagagagcgga	4174
QY	4144	tcacctcgtcagacttccatccctcgtgtgcaaggttcaatgttaagaaagcgaaaccgcgc	4203
Db	4175	taccactccctgggtctctcttccacagcttcaactcttcaacttgaagaaagcggtatccacgc	4234
QY	4204	caatttcgagcaagaccttatacaagccgggcatttcgcgcgttggacagacttggcgcg	4263
Db	4235	ctgtgtctgagcaagcgtctgtacacggcgtgcacattccacbtccgataataatcaacaggg	4294
QY	4264	aattgtctacgttcaagggcgagccacacagaagaagacacatcaactcaatcaacatagacc	4323
Db	4295	aactactacccgttcgtgcgaactcatctcgaataatccgaataatgacatatatacatcgag	4354
QY	4324	gtgcgagcatcagctcgtgacagcagccttaaaacggtgtgcgagacttcogtccgtcgcgcgc	4383
Db	4355	acggtctctatgtgcggtgagactccaactgtgaaacggtcaagagacttcgggtgttccatcga	4414
QY	4384	atgcgacacacggcggtcgttctgcgttcaataatggagcccccgccttccatgcagcgatgt	4443
Db	4415	acgcgcgaacggcggtctcctcaatactatgtggatataacactactcgtccagcatgcagcatgt	4474
QY	4444	tcgagttcgacgtcatcgcgtctacgataacagcttcgcgaatccgaatccagacccgtgtgaaag	4503
Db	4475	ttgagttcaacgtcatcgtctacttaccaccaatgtagaagacgaataagcgagaggttgaaag	4534
QY	4504	tctactcatctcgtgaaaaacgggtgactcttcaatttgcataaaccaacttgcagac	4563
Db	4535	tctactcatcttcatcccaaaatagggtgtgccttcatatctcctgaaacgagtgtggagacgg	4594
QY	4564	tttgacgaacagaagaatttctatagcgcacagtttcaagcacccgggttccaatgtagtgcga	4623

Db	1391	gaggagagccttcctccacatcaacaaacgtgtgtcattcattgtgacagacatcacaaccca	1450
Oy	1326	cgaccgcgcgcgttcataaagaatacagtatgtatcatcattgagagaactccaatgtact	1385
Db	1451	acacctgaacctatacacaaagaataatcgcactgcggaatcatgtgagagcgccctgacc	1510
Oy	1386	ctaaactcaatlgaaaatcttgatctcatatgcagatctgtgtgtgaaacgtcataatc	1445
Db	1511	ctcaactctccataaagaattcggattctcatgtataagattttagatgctaaacgtcagtaac	1570
Oy	1446	acagtgagaaacttgagagacgtgtctccgcgaaggcggtccgcattctcatctgcgcgcg	1505
Db	1571	acgggtgcgtcttagagagagcgttgtaacctccacccgctgtcgtgagcatctcataatgacct	1630
Oy	1506	gggagcggcttaccgaagagcagaaccttcacatcaltgtgtgtgcaacataaacacacatgtctgat	1565
Db	1631	gaagtcgcgtactacagcagacagacacttcacatcgtccacacctcattcacatcactgtgtgat	1690
Oy	1566	tacgaagaatgttactttttttagaacaatcatcatctaaggtgtcaaaagcagtgagatlgaaacaac	1625
D	1691	tacgaagaatgtccagagtttcaagatataagatattcgggtgtgttagtgacgcgaacaacaagac	1750
Oy	1626	gctaaccaactgttgccgaagcgcgtcgtgtgtacgtgtgaaccttcataactcgtgaacgcgaactc	1685
Db	1751	acgagagcaagttgggtgcgttcgctgtgtgttcaactatccactcatcaattgtgaacgagttgcag	1810
Oy	1686	cccatcttcgaagagagacagactactccgcgttcgtttaaagagacggtctggcgcgcgtcttc	1745
Db	1811	ccgattcttcgaacaacgcgtgtgacagacgttcacacttcgcagcagactcaaacgcgaggggttc	1870
Oy	1746	ccgggtggccacagtgctgcgcctccgcgaacagagaatcagacagactaagtgatctaatcgt	1805
Db	1871	tccgttcgcgaagcgggtgtgcacagcagagaaatcgtggagtgctgtcgcagcatacttta	1930
Oy	1806	atgggcaacgcgtgtgtgactactcgttcatagatgaatcaacgggagagatcttcgtgagc	1865
Db	1931	tgtgggtaaacgtgtttaaacttctcgtacatccatcgaacaactccacgcgcgcgtctca	1990
Oy	1866	atggaagagactccttcgaactacacacgcgacagaaacactatcttgtagtgcgcgtctgac	1925
Db	1991	gtctaacgaactccttcaactacacacacgcgagaaagtgaattatttgtgcaggtgcgcgtctaca	2050
Oy	1926	gactactttgggcgaacggcccaacaaacagtgacacacccagctgtgtgatagaactggag	1985
Db	2051	gacacgcgtgggga---acaccttcacaacgcgcgaacgttcacagctgtgtcatatcagactaat	2107
Oy	1986	gatgtataaacaacactctctccacacccatcaggtgtcccgcttcgcactcccaagcgttcgaagag	2045
Db	2108	gacataaacaacacgcgcacccacacttaccgtgtccttcgaagcagttccccaagtgtgaagag	2167
Oy	2046	aacgttcccgaaagatacagagatactccgcgggaatcatctgtacacgcgaccccgacacagc	2105
Db	2168	aacgtgtcgtgatgtgcgaacgtcatcacaccagaggtttagcgcgaacgcgcacccgacaacag	2227
Oy	2106	gectactcgttgttcgaagatcgcactgtggactccactcgtggccacacaaagcgagagag	2165
Db	2228	ggcgaactcgcgtctcgaataaactgtggaacactcttctgcacacaagcgaagcgccag	2287
Oy	2166	accaaacacttgaataacgtcgcgtgtatagttatcgaataaagatataacccacacgaagggc	2225
Db	2288	gtcaaaccccgacgaagtttagaattgcgtggaatctgagacacatcttcccgagattaac	2347
Oy	2226	aaccgcgggttcgcgcacatcgcgcgcctctcgtgtgcaagagatccgggaacaagttcacatc	2285
Db	2348	aaccgcgggactgtgtatcgcgcggtgtgtagcgcggaatcagacaacaaagtgacata	2407
Oy	2286	gacttcgaggaatttcgagatgcttactactaacgctccgcgtgtgagggactcccaaacgttc	2345
Db	2408	gactacgaagagattttaggtgtccctccctcacaagtgaggtgtgcgttcaaacgctc	2467
Oy	2346	atcggaatctactacgattgagcgcgttcaagatccacaataatcgcacatgaaagacaac	2405

[illegible]

XX WPI; 2001-329223/34.
XX
XX New nucleic acid encoding Bt toxin receptor, useful for screening and
PT identifying new Bt toxin receptor ligands useful as new insecticidal
PT toxins -
XX
PS Example 1; Page 83; 85pp; English.
XX
CC The present sequence of is that of a partial cDNA clone obtained
CC by RT-PCR amplification of European corn borer (ECB), *Ostrinia*
CC nubilalis, larva midgut cDNA. The clone was used to screen a
CC midgut cDNA to isolate further clones, from which a full-length
CC sequence (see AAF30933) encoding the Bacillus thuringiensis (Bt)
CC toxin receptor (see AAB20498) of ECB was deduced. The polypeptide
CC encoded by the present clone was also used to inoculate rabbits for
CC polyclonal antibody production. The ECB Bt toxin receptor binds
CC CryIA, particularly CryIA(b), toxins. The invention provides
CC compositions and methods for modulating susceptibility of a cell
CC to Bt toxins. The compositions include Bt toxin receptor proteins
CC from the lepidopteran insects ECB, corn earworm (*Heliothis zea*) and
CC all armyworm (Spodoptera frugiperda). The methods are useful for
CC investigating the structure-function relationships of Bt toxin
CC receptors, investigating toxin-receptor interactions, elucidating
CC the mode of action of Bt toxins, screening and identifying novel
CC Bt toxin receptor ligands including novel insecticidal toxins, and
CC designing and developing novel Bt toxin receptor ligands. The
CC methods are useful for managing Bt toxin resistance in plant pests,
CC and protecting plants against damage by plant pests.
XX
XX Sequence 1604 BP; 386 A; 465 C; 419 G; 334 T; 0 other;
SQ
Query Match 28.1%; Score 1547.4; DB 22; Length 1604;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 3040 tcttgaataatcgtatgaatcttcttgaatgagtgaggaaccagagcg 3099
DB 8 tcttctcaactcagacacactcttcttgcagtgagcgtaggagaaccagcg 67
QY 3100 aagtgtaaatattgtcgtcttattgtgatgtgaacgaacagctccctgaatgcattgc 3159
DB 68 aagtgtaaatattgtcgtcttattgtgatgtgaacgaacagctccctgaatgcattgc 127
QY 3160 ctgtgtaactcgtgttgaatttcccgaaagagcagtgctgctgcttaccctccag 3219
DB 128 ctgtgtaactcgtgttgaatttcccgaaagagcagtgctgctgcttaccctccag 187
Q 220 aaatctacgacccgagcagagatgaacacgaacagcagcagctgctgcttaccgaa 3279
DB 188 aaatctacgacccgagcagagatgaacacgaacagcagcagctgctgcttaccgaa 247
QY 3280 tctcggaactcagatcagcagcagacatcgagtgctgagatctcttaccatgatct 3339
DB 248 tctcggaactcagatcagcagcagacatcgagtgctgagatctcttaccatgatct 307
QY 3340 cgattgaaacaaactgggaacttgagccgctatgagattgaggggagattgggca 3399
DB 308 cgattgaaacaaactgggaacttgagccgctatgagattgaggggagattgggca 367
QY 3400 cttaacgaatactcatgaagccttcgacacgagctacccgagagaggtccaaacgga 3459
DB 368 cttaacgaatactcatgaagccttcgacacgagctacccgagagaggtccaaacgga 427
QY 3460 cgtacacccgtgcatccgccttacaacttccacacccgtgtgtcgttcccgcaac 3519
DB 428 cgtacacccgtgcatccgccttacaacttccacacccgtgtgtcgttcccgcaac 487
QY 3520 ccgactccgctatctcgtgttctagggagcgagcaagaagcgcggtctctggcagag 3579
DB 488 ccgactccgctatctcgtgttctagggagcgagcaagaagcgcggtctctggcagag 547

QY 3580 ctgcaacagatctctgtagccgatactacgacacgagacagcgctccacgagcgca 3639
DB 548 ctgcaacagatctctgtagccgatactacgacacgagacagcgctccacgagcgca 607
QY 3640 gcttcacgttccacgctccaggaagaatgagagcgcttcagtaattgataactgaag 3699
DB 608 gcttcacgttccacgctccaggaagaatgagagcgcttcagtaattgataactgaag 667
QY 3700 tgggagcaggaagaatgagagcgcttatactacgacacgcttccacgaagaatca 3759
DB 668 tgggagcaggaagaatgagagcgcttatactacgacacgcttccacgaagaatca 727
QY 3760 gacattcagatcagatcccgagccacgagcagcgagcagccgcccgtttgga 3819
DB 728 gacattcagatcagatcccgagccacgagcagcgagcagccgcccgtttgga 787
QY 3820 ccgagctacgttcttggtgtgtcttctgtaaccacagagggcgaccgtgttcagcga 3879
DB 788 ccgagctacgttcttggtgtgtcttctgtaaccacagagggcgaccgtgttcagcga 847
QY 3880 atgcagctactgtgcctctcttcgaggtggaagaagcctcgtgagagtttgaagctgc 3939
DB 848 atgcagctactgtgcctctcttcgaggtggaagaagcctcgtgagagtttgaagctgc 907
QY 3940 cgcagcagaaagacattaaaaacacactctgcgaagaatgcagtgccaaagatatcactaca 3999
DB 908 cgcagcagaaagacattaaaaacacactctgcgaagaatgcagtgccaaagatatcactaca 967
QY 4000 ggttattgagcgaacaaagaaggtcttctgtaactgagcagatcaagcagctcatct 4059
DB 968 ggttattgagcgaacaaagaaggtcttctgtaactgagcagatcaagcagctcatct 1027
QY 4060 cccctgcagagatgttgacccgagagtgagcagcgtcttacaagctgcagcgca 4119
DB 1028 cccctgcagagatgttgacccgagagtgagcagcgtcttacaagctgcagcgcgca 1087
QY 4120 gaaactgcgcgagcgaactggagatccctctgcagacttccatccctggttcaagctca 4179
DB 1088 gaaactgcgcgagcgaactggagatccctctgcagacttccatccctggttcaagctca 1147
QY 4180 atgtaagaagaagcgaacccgcgccaatttccgaagaagacttcaacacagcgagctt 4239
DB 1148 atgtaagaagaagcgaacccgcgccaatttccgaagaagacttcaacacagcgagctt 1207
QY 4240 cgaagcttgagacagcattggccgagatgtctactgttcaaggcgagccacagaagaagcg 4299
DB 1208 cgaagcttgagacagcattggccgagatgtctactgttcaaggcgagccacagaagaagcg 1267
QY 4300 aacacatcagtaacacatagacgctgcgagcatgagcttgagacagcagcctagaagcgcg 4359
DB 1268 aacacatcagtaacacatagacgctgcgagcatgagcttgagacagcagcctagaagcgcg 1327
QY 4360 tgcgagactcgccttcgctgcgtgcgtacgtgacacacgagcggtcttgcgtcaatgagc 4419
DB 1328 tgcgagactcgccttcgctgcgtgcgtacgtgacacacgagcggtcttgcgtcaatgagc 4479
QY 4420 ccaacgcttcatgacagcagctgttcgagcttcgagctcatgcgttaccagatacagcttcg 4479
DB 1388 ccaacgcttcatgacagcagctgttcgagcttcgagctcatgcgttaccagatacagcttcg 1447
QY 4480 caatcagacagcccggtgtgaaggtctactcctcatctcatcgaacaaacgagtgacactca 4539
DB 1448 caatcagacagcccggtgtgaaggtctactcctcatctcatcgaacaaacgagtgacactca 1507
QY 4540 ttctcgataacacacttgagacgcttgagcagagaagaatttcatagcgcgcaagcttca 4599
DB 1508 ttctcgataacacacttgagacgcttgagcagagaagaatttcatagcgcgcaagcttca 1567
QY 4600 gcaacgggttcaacatgacgctgcaacatcgagccaggt 4636
DB 1568 gcaacgggttcaacatgacgctgcaacatcgagccaggt 1604

```

RESULT 7
AAF30935
ID AAF30935 standard; CDNA; 5592 BP.
XX
XX AAF30935;
AC
XX
XX 09-JUL-2001 (first entry)
DE
XX Spodoptera frugiperda Bt toxin receptor CDNA.
XX
XX Fall armyworm; FAW; Bacillus thuringiensis; Bt toxin;
XX receptor; crystal protein; CryIA; biological control; insecticide;
XX crop protection; ss.
XX
XX Spodoptera frugiperda.
XX
XX Key Location/Qualifiers
XX CDS 162..5366
XX sig_peptide 162..227 /*tag= a
XX mat_peptide 228..5363 /*tag= b
XX /*tag= c
XX
XX WO200136639-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-US31674.
XX
XX 18-NOV-1999; 99US-0166285.
XX
XX 21-SEP-2000; 2000US-0234099.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Flannagan RD, Mathis JP, Meyer TE;
XX
XX WPI: 2001-329223/34.
XX
XX P-PSDB; AAB30500.
XX
XX New nucleic acid encoding Bt toxin receptor, useful for screening and
XX identifying new Bt toxin receptor ligands useful as new insecticidal
XX toxins -
XX
XX Claim 1(a): Page 72-79; 85pp; English.
XX
XX The present sequence of is that of cDNA encoding the Bacillus
XX thuringiensis (Be) toxin receptor (see AAB20500) of the fall
XX armyworm (FAW), Spodoptera frugiperda. The receptor protein binds
XX CryIA toxins, particularly CryIA(b) toxins. The invention provides
XX compositions and methods for modulating susceptibility of a cell
XX to Bt toxins. The compositions include Bt toxin receptor proteins
XX from the Lepidopteran insects FAW, European corn borer (osteia
XX nubialis) and European corn borer (Heliothis zea). Nucleic acids
XX encoding the polypeptides, antibodies specific to the polypeptides,
XX and nucleic acid constructs for expressing the polypeptides in cells
XX are also provided. The methods of the invention are useful for
XX investigating the structure-function relationships of Bt toxin
XX receptors, investigating toxin-receptor interactions, elucidating
XX the mode of action of Bt toxins, screening and identifying novel Bt
XX toxin receptor ligands including novel insecticidal toxins, and
XX designing and developing novel Bt toxin receptor ligands. The
XX methods are useful for managing Bt toxin resistance in plant pests,
XX and protecting plants against damage by plant pests.
XX
XX Sequence 5592 BP; 1662 A; 1301 C; 1332 G; 1297 T; 0 other;

```

```

Query Match 26.7%; Score 1465.8; DB 22; Length 5592;
Best Local Similarity 58.0%; Pred. No. 3.7e-312;
Matches 269; Conservative 0; Mismatches 2037; Indels 111; Gaps 17;
Oy 160 aatggtggttgagaggtcttcacagcagtgctactgtctcttaagcctcgcgcac 219

```

```

Db 160 agatggtggttgagaggtcttcacagcagtgctactgtctcttaagcctcgcgcac 219
Oy 220 tagccaaccaagatggtctgtacattatcgcataaccagacggagactccggactgc 279
Db 220 cacagcgaagatggtggtctgtatgataaccagacgagcgtgactcc 279
Oy 280 cgcctatgattacgaagaataatcagtgagtgaaacgctctataccgcccagacc 339
Db 280 caactcaaatcttgacggtttacaatggttcagagccactattaccagtgagacc 339
Oy 340 gagaggaatgtagtggaaactcttcacgatacaatga-----ttc 384
Db 340 gagaagaggtcgcctccatcgaatgaaacccgataccctggagcaacaacatgtagcc 399
Oy 385 aggtcatatatacaggaagaagaatcgaagaagacgtcatcttcggaagcttaactac 444
Db 400 agagaattacatgtagggagagatcgaaggtccgtagcatctgcgaataataactacc 459
Oy 445 aaggttccaacacgcccgtgctgctgattatgtcaagccagccagagccca---gtcgg 501
Db 460 aaggaacaacccccctcctcaaaataagattaccttctgtgtgtgcagccacatgctg 519
Oy 502 gccctgagttcg---acagaatgaagcagacggccaatgagccttgattacgcaaa 558
Db 520 gagcagaagaattcgtgaataatccttgacgcgaactgagagctgtgatttgaataactcaa 579
Oy 559 gacaagactacgagacgaacgaacatcgcagagctatggttctcaatccaagtggaggtg 618
Db 580 ggcagagcatatgaactctctgtatgcagagatacagcttgcagtgtggtggagagcc 639
Oy 619 aatcacagcgctactgtgtggtggtgagatgtaacatcgaacgaacatccgcccacc 678
Db 640 agtcctgtgttgtaacggtgagagctggatgtgtaacatcgaacgaacatcgcgcacca 699
Oy 679 tgcaggtgtcagcgctgcgtgaattccagaacatcggcaggtcgtgactgcagcactgcg 738
Db 700 ttgagatgttgagccttgcgaacttaccggaactgttgaacccatgttaacagatgta 759
Oy 739 tgtaccaagtgtcagacgcgcgcagtgtaaatcagacccgccttactgacgttcgtgc 798
Db 760 aataatacgtgctcgcagcgaacggtcgtgatcagtaacagtgttatgtatcatatag 819
Oy 799 acagacgaaggtctgcagatgtaagaatccttcatatggttggagaataacgcccagcg 858
Db 820 acagcgaag---aggaaacgaagaagaatctcgaactgacgaagaagatataccggcg 876
Oy 859 actggtcaataatgaagaatgacttggggatcaatcgccttgcactgcagacaactc 918
Db 877 attgacgaaggtgtatattgtcttgtaattgaanaaatcctcttgattagagaagaatc 936
Oy 919 agcttcataatttagcgttcacagcttctgcactgcctcgcgaacaacacacagtgtaaca 978
Db 937 ctctacacatattcagatcagctgttctgtatcttccctcaacaacatgagcgtgtgca 996
Oy 979 tgatgtgcaagtgaagaacgtagaggtctgcgcccctcgcgtgggttgagatcttctag 1038
Db 997 tgaatgtggaatgaagaacgtagaacaatagaataatccctcgtgtgagatcttctgctg 1056
Oy 1039 tgcagcagtttgacagagaactaactcagagcttctcctccgcgcgtatacggggagca 1098
Db 1057 tgcacagtttgatgaanaaacaagcggaatcgttcaacagtgtagcgtatttgatgtagca 1116
Oy 1099 cgggaatcaatagggcactcaactataccctcatcagagatgagcgtcagcaactctttt 1158
Db 1117 cgggaatcaatcaactatctatcgtlatagaagaatgaagaatgaagaagaagtctt 1176
Oy 1159 -----ccctgaggttgatgaagcagagcattctcgtgactgtgactgagcgtg 1206
Db 1177 tcaagcattgagaacaataggggaaagcagagcagcgtgacagatccacgtgctccctatag 1236
Oy 1207 accgcgaagaagcttgaagaagagcttttcaacctcaccatcgttgc----- 1252

```

Db 1237 acagagactacctgaanaaggataltgttcataataaatatgtacatatataaagaagtgt 1296
OY 1253 -----ttacaatatctactagcgttagctttgtcaacagagccacatttcacatcgttca 1308
Db 1297 ataatagaacaagaaggtgtaatacatcgttcgagacctcagcaaatgtgacgtatataalta 1356
OY 1309 acgacgtlcaatgatcagagacccgagccgtgtcgtataaagaatatagatltgtatcatgtg 1368
Db 1357 acgatataatgatcagagccgagaccccttcataaagaatatcacgtatcttcataatgtg 1416
OY 1369 aggaaccccaatgactcttaactcaatgtgaagaatttggatccatgtatcgtgaagtttgg 1428
Db 1417 aagaacgtcgatgaccttaatttgg-----caagagtttggtttccatgtacgttgcattgt 1473
OY 1429 gtgaacacgtccataacacatgtgagacttggagacgtgttcccgccagggcggtgtccg 1488
Db 1474 gtccccaacgctccagtaacgacgttcccttagagagatataacagccaagggggccctaccg 1533
OY 1489 catlctacatcgcgcgcgggagcggtacacagagcagaccttcaatgtggtccacatataa 1548
Db 1534 cttctacatcgcgcgcctgaagaaggtttacacagccagcttcttcacataagttactagaa 1593
OY 1549 accaacaatgcttgatattacgaagaatgtcatatt-----ttcagaacaatcatatgaagtca 1605
Db 1594 tccataaacaatgttgaattatagaagaatgacgactacagaccaggaataaagctaaagcgag 1653
OY 1606 aagcaggtgagactgacacacagcttagacacgctgtgacgagcgctgtgtatcgtgaacgtga 1665
Db 1654 tagcattgtgacagacagcatataacatacatgtgggaagcaattttaacatataactta 1713
OY 1666 tcaactgtgaacgacgaaacttcccatcttcgagagagacgactactcgcgttcttaagg 1725
Db 1714 tcaattgtgaatgatgacttaactataatcttcgacgagacgcttaacaacgttgcatttggg 1773
OY 1726 agacgctgtgcgcgcgtctcccggtgtgcacagcggtgtcgccttcgacagagacatcgagc 1785
Db 1774 agacgctgtgcgtgtatgtctccacattgtglaaataccggtctaaagaacagagacatcggtg 1833
OY 1786 agctgtgattgtgattcatgtgatgtgggcaacgctgtgtgactactgttctatagatgata 1845
Db 1834 acatgtctgagacactcgtatatgtggcaacgctgtcaaaacttccctgagaattgtcacataga 1893
OY 1846 cgggagagatctctgtgagcatgtgacgtatgacctgacacacagcagaaacactctat 1905
Db 1894 ctggagatgtgtaacgtgtgtcaacgagacgacttacttgattatcaaaagacgaacgaatca 1953
OY 1906 ttgtctcaggtgcgcgctcgtacgatacttgggcgcgacggtccacacaaacagatgtcacccc 1965
Db 1954 tagttcaaatctgtgctgttgaatacactagg-----ttaacctgaagaacaggggttaacacac 2010
OY 1966 agctgtgtatgaactgtgagagatgtcaacaacgactctccccccttacgcttgcgcggt 2025
Db 2011 agctcaagatatatttttgaagaagacatacaacacgacacacttaactgtcagctgcaggt 2070
OY 2026 cgaactccaagcgttcgagagaaagcgttcccgaaaggaatacagagataatcccgggaaactactg 2085
Db 2071 ccagtcacaagtgtgaagaagacgttgaagtcgggacacccgttaacgaggggtcaacgg 2130
OY 2086 ctacgcgacccgagacacacgacgctcactcgtgtgttcgagatcgcactgvggactccactgg 2145
Db 2131 cgaacgaccccaagacacacacacgacttatacacttgcagatcgtatggtggaaacttcttgcg 2190
OY 2146 ccacacaaagaggg-----cagagagacacaaacctactgaaatagtcgggtgtatagttacg 2202
Db 2191 ctacgaagacgagggccacaaatgtgacccaacacactgtcagactacacagatgtcgtagaatcc 2250
OY 2203 aaacgatatatacccaacgagggcaacgggtgtccgcacatctcggtgcgcgtccctgtgtgtcaag 2262
Db 2251 tgaacgtatatacccgatctcgtacaatacgaaggagagcgttgaagggtcaactgtgtgtgacggtg 2310
OY 2263 agatccgggacaacgctcaacatcgaactcgcaggaattcgcagatgtcttaactccacgctcc 2322
Db 2311 aggtcagatgtgtgcgtgacacatcgtatcagagaagtttgaaggtgtgtgtaacctcgtctgtca 2370

OY 2323 ggtgtgaggaactccaacactgtlcatcggagatgaactacgaatgagggcagcttcaacatca 2382
Db 2371 ggtgtgatatgacgaacactgtlcatctgtgccttgattatgacgaagcaatgtctgacgggtga 2430
OY 2383 caataatcgaatcgtgaacgaacacggtccgactcttcgcgaacggcagcgtctgacgagatga 2442
Db 2431 cgaatatcgaatatgaacgaacacttgcgatatgtggccgaacacacgcttgcagagataac 2490
OY 2443 tgcgcgtgcgcgagctgtgcgcgcacgacgacgtcactcgtgcctgcgtgtcgcacagaca 2502
Db 2491 tgcgcgtgcgcgagatgtgcgcgcagcagcgtcactcgtgcctgcgtgtcgcacagaca 2550
OY 2503 tgcagcgcgcgcgtctacacacgaagtgtgcgttcaactataacacttagaanaacacactccg 2562
Db 2551 tgcagcgcgcctctcaacacgagtcgcgttcaacacactgtgtccccaatgaagacactcgt 2610
OY 2563 aggaattagtgaaagttaacttcaacactgttcaaatltgaagttgatatcgtcgaagggcga 2622
Db 2611 atgacctataagcgaatcactacgtacacgttcaacgttgcactgtgaacaaagggcgaagca 2670
OY 2623 tgcagtcgaacgaacccgtgcgcgtcttacttgaactacacgttcaactcgtcgtacgagat 2682
Db 2671 tgcagcgcgaatgatccacactcgtcttactcgttatcaaaagttcaacttgcgcgataagt 2730
OY 2683 gctccctgtgaanaa-----ccgcacaggaatgtlccctcagatctccaactactcgaagttccag 2739
Db 2731 gctcctcgtacgagttcttccctgtgtgcaccaactgtgaccccaacttactcgtgaataccagg 2790
OY 2740 ggcataatcgaataataatcatcgtacgaacacaaacaaagtgtcgttgaacccgtctcatatga 2799
Db 2791 gagagatagcgaatcgcgtataacggaatacgaacaaatltccacgcgcgtggaanaaata 2850
OY 2800 agttcaacaacgcggtgtgaagctcgtgaggaatgacacgagcgcgcgtgacgaaggtgtccagc 2859
Db 2851 tgtlccctagtgaaagcgcacatctatgtgaacacacccaaatgttaccagaatcacagaca 2910
OY 2860 tgaactccacagaccgttgacagagacgagttgttaccacaggtltagacatacgaatgaact 2919
Db 2911 tcatcgtcagtgatgacacagagatgatacgaacaaataaagcgtgtgcatacgaatcaact 2970
OY 2920 ttgcgcgtgaaccccgactgcgggaattcttcgcgaagtgagacctgtgacacatgtgccttg 2979
Db 2971 acgcattcaacacacagcgtgtgaagaactcttcgcagttgacacctgatatatcgttgaactgt 3030
OY 2980 aggtgcatltaacccggggagacgaanaaatgtgacgcgactgtgggagatgaagcttaacata 3039
Db 3031 ttgtccacttccacactagcgaagtgtgtgacagagacggagagaaacggagacataaga 3090
OY 3040 tcttctgaatttcaatcgaataactcttcttctgagtgagcgttgaaggtgagaaacggagacg 3099
Db 3091 tcatcttcaacactcgtcgtatgaatcgtggaagcgtcgtggagatgtgcaatcaagaacacatct 3150
OY 3100 aagttgaatatltgtcgttctatgtgatgtgaaacgaacgctcctcgtatgacatgtgc 3159
Db 3151 ccacgagaggtgcgtgttatactcgttgaataaagacaaataaagccgggaactacaacttcc 3210
OY 3160 ctgaggaactccggttltgatgttctcgaagggcaggttgcgtgtgtcgtgtgactccag 3219
Db 3211 ctgagggcgaatttggacgcgttctccgaagggtgaaggtggaagaaacgcattccaacag 3270
OY 3220 aaatcagcagccggaggtgagtgaaacgagacgaacaaactcgtgtgtlccgttgcggaa 3279
Db 3271 agatcagcagacacagacagagatgaacatcaacaaacactcgtcgttggatatgtgaa 3330
OY 3280 tctgtgacctacgaatcagcagcagacacatcgaagttgc-----agatctcttccacatga 3336
Db 3331 ttcgattcgaataattgtatcaatagaacatcgaactcctcctaagaatcaataataa 3390
OY 3337 tctcgaattga-----aaacaaactgggggaacttgaagccgttatgagact 3381
Db 3391 taacgattgagatcgtcgtatcactcgttgaanaattcgtgtgagagtttgaagactacacatgagcc 3450

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 13:48:22 ; Search time 74.72 Seconds

(without alignments)
1702.139 Million cell updates/sec

Title: US-09-715-909-2

Perfect score: 8909
Sequence: I-MGVERFPAVLVLVSLASAL.....THGNNGFKSTPYLPQHPK 1717

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq-1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8909	100.0	1717	22	AAAB20498
2	5178.5	88.4	1730	22	AAAB20499
3	4923	55.3	1729	22	AAAB2331
4	4693	52.7	1528	20	AAAB20182
5	4561.5	51.7	1734	22	AAAB20500
6	4408	49.5	1528	17	AAAB5333
7	519	5.8	2923	22	AAU07054
8	519	5.8	2956	22	AAU07053
9	494.5	5.6	3298	22	AAE03657
10	459.5	5.2	2707	18	AAW27161
11	444.5	5.0	3014	22	AAU02196

12	324	3.6	1026	15	AAAB58906	Human protocadherin
13	324	3.6	1026	17	AAAB7146	Protocadherin clon
14	324	3.6	1203	15	AAAB58911	Product of alterna
15	324	3.6	1203	17	AAAB7152	Alternatively splic
16	317	3.6	2405	21	AAAB42192	Human ORFX ORF1956
17	311	3.5	1180	19	AAAB61273	Homo sapiens proto
18	305	3.4	788	22	AAAB48370	Human SBC3 protein
19	303	3.4	916	20	AAAY41743	Human PRO707 prote
20	303	3.4	916	21	AAAB44299	Human PRO707 (UNO3
21	300	3.4	627	21	AAAB38349	Human secreted pro
22	296	3.3	896	19	AAAB41252	Xenopus paraxial p
23	295.5	3.3	797	17	AAAB6866	Human protocadheri
24	291.5	3.3	882	20	AAAY09375	Wild-type E-cadher
25	291.5	3.3	882	21	AAAB35730	Human E-cadherin a
26	291.5	3.3	882	22	AAAB73490	Human E-cadherin
27	288	3.2	100	22	AAAB82332	Plink bollworm toxi
28	285	3.2	558	22	AAAB38969	Human polyepitide
29	284	3.2	772	20	AAAY41725	Human PRO941 prote
30	284	3.2	772	21	AAAB44281	Human PRO941 (UNO4
31	284	3.2	772	22	AAAB31209	Amino acid sequenc
32	283.5	3.2	889	20	AAAY24913	Human ontherin. H
33	283.5	3.2	889	20	AAAY21687	Cadherin-like poly
34	282.5	3.2	1184	20	AAAY41750	Human PRO731 prote
35	282.5	3.2	1184	21	AAAB44306	Human PRO731 (UNO3
36	282.5	3.2	1184	22	AAAU12368	Human PRO731 polyp
37	282	3.2	1187	22	AAAF06028	Human asthma-assoc
38	282	3.2	1187	22	AAAB56684	Human protein sequ
39	280	3.1	831	22	AAAB31106	Peptide #5143 enco
40	277	3.1	800	21	AAAB18920	A novel polypeptid
41	274.5	3.1	653	18	AAAB56557	Rat cadherin-10.
42	274.5	3.1	653	18	AAAB31128	Full length rat ca
43	274.5	3.1	878	15	AAAB55060	Sequence of human
44	274.5	3.1	878	16	AAAB5487	Human E-cadherin p
45	274	3.1	906	21	AAAY0741	Human N-cadherin.

ALIGNMENTS

RESULT 1	AAAB20498	standard; Protein: 1717 AA.
ID	AAAB20498;	
AC	09-JUL-2001	(first entry)
DT		
XX		
DE	Ostrinia nubilalis Bt toxin receptor.	
XX		
KW	European corn borer; ECB; Bacillus thuringiensis; Bt toxin;	
KW	receptor; crystal protein; CryIA; biological control; insecticide;	
KW	crop protection.	
XX		
OS	Ostrinia nubilalis.	
XX		
FT	Key	Location/Qualifiers
FT	Binding-site	1293..1462
FT	/label=CryIA_binding_site	
FT	Domain	1571..1589
FT	/label=Transmembrane_domain	
XX		
PN	WO200136639-A2.	
XX		
PD	25-MAY-2001.	
XX		
PF	17-NOV-2000; 2000WO-US31674.	
XX		
PR	18-NOV-1999; 99US-0166285.	
XX		
PR	21-SEP-2000; 2000US-0234099.	
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Flannagan RD, Mathis JP, Meyer TE;	

XX MPI: 2001-329223/34.
DR N-PSDB: AAF30933.
XX New nucleic acid encoding Bt toxin receptor, useful for screening and
PT identifying new Bt toxin receptor ligands useful as new insecticidal
PT toxins.
XX Claim 4(a): Page 57-61; 85pp; English.
XX The present sequence of is that of the *Bacillus thuringiensis* (Bt)
CC toxin receptor of the European corn borer (ECB), *Ostrinia nubilalis*.
CC The sequence was deduced from cDNA isolated from an ECB larva
CC midgut cDNA. The receptor protein binds CryIA toxins, particularly
CC CryIA(b) toxins. The invention provides compositions and methods
CC for modulating the susceptibility of a cell to Bt toxins. The
CC compositions include Bt toxin receptor proteins from the
CC lepidopteran insects ECB, corn earworm (*Heliothis zea*) and fall
CC armyworm (*Spodoptera frugiperda*). Nucleic acids encoding the
CC polypeptides, antibodies specific to the polypeptides, and nucleic
CC acid constructs for expressing the polypeptides in cells are also
CC provided. The methods of the invention are useful for investigating
CC the structure-function relationships of Bt toxin receptors,
CC investigating toxin-receptor interactions, elucidating the mode of
CC action of Bt toxins, screening and identifying novel Bt toxin
CC receptor ligands including novel insecticidal toxins, and designing
CC and developing novel Bt toxin receptor ligands. The methods are
CC useful for managing Bt toxin resistance in plant pests, and
CC protecting plants against damage by plant pests.
XX Sequence 1717 AA:
SQ
Query Match 100.0%; Score 8909; DB 22; Length 1717;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVERFAVLVLSAALANORCSYIIAPREPPELLPIDYEGKWSQPLIGPTR 60
DB 1 mgverfavlvlvsaaalancrcsyiaiprepellpidyegkwsqpligptr 60
QY 61 EEVCKENLPROMIOYITVEEIEEDYIAKLNQGSNTPLVSTMSGPRQAGFEFRON 120
DB 61 eevckenlprpmioyitveeieedyiaklnqgsntplvstmsgprqagfefron 120
QY 61 eevckenlprpmioyitveeieedyiaklnqgsntplvstmsgprqagfefron 120
DB 61 eevckenlprpmioyitveeieedyiaklnqgsntplvstmsgprqagfefron 120
QY 121 EADGWSLVITROOYEFATMOSYVESIOVEGSOAVLAETIVINDINPILQVVSACV 180
DB 121 eadgwsvlvitrooyefatmosyvesiovegsoavlaetivindinpilqvvsacv 180
QY 181 IPEHGEARLTCVCYVSDRDGEISTRFMTFRVDSGRADESIFYMVGEYDPSDMFNKMT 240
DB 181 ipehgearltcvcyvsdrdgeistrfmtfrvdsgradesifymvgeydpdmfnkmt 240
QY 241 VGINSPLNFETTQHLIFSVYASDSLPPNNHTVTMVOVENESRPPRWYEIFSVOGFDEKT 300
DB 241 vginsplnfettqhlifsvyasdslppnnhtvtmvovenesrpprwyEIFSVOGFDEKT 300
QY 241 vginsplnfettqhlifsvyasdslppnnhtvtmvovenesrpprwyEIFSVOGFDEKT 300
DB 241 vginsplnfettqhlifsvyasdslppnnhtvtmvovenesrpprwyEIFSVOGFDEKT 300
QY 301 NOSESLRAIDGTGIRNAINNTLIRDDADDFSLVIEGALHTETIDROKLEBELFNL 360
DB 301 nose斯拉IDGTGIRNAINNTLIRDDADDFSLVIEGALHTETIDROKLEBELFNL 360
QY 361 TTVAAKSTPDASATPAHTFIIVNDVNDGRPEPLHKEYSIDIMEETPMTLNFENBERGFHDR 420
DB 361 ttvAAKSTPDASATPAHTFIIVNDVNDGRPEPLHKEYSIDIMEETPMTLNFENBERGFHDR 420
QY 421 DLGENAQYVELEDVFPFGAASAFYIAPGSYQROTFLMGITNHTMLDYEDYIFONIIK 480
DB 421 dlgenAQYVELEDVFPFGAASAFYIAPGSYQROTFLMGITNHTMLDYEDYIFONIIK 480
QY 481 VAAVDMNNAHSHGEALYVNLINMDELPIFEBSYSAFRTYVAGCFVAVVLAIDRI 540
DB 481 vAAVDMNNAHSHGEALYVNLINMDELPIFEBSYSAFRTYVAGCFVAVVLAIDRI 540
QY 481 vAAVDMNNAHSHGEALYVNLINMDELPIFEBSYSAFRTYVAGCFVAVVLAIDRI 540
DB 481 vAAVDMNNAHSHGEALYVNLINMDELPIFEBSYSAFRTYVAGCFVAVVLAIDRI 540

QY 541 DRYVYHSLMGNAVDYLFIDESTGEIIFYSMDADFYHRONTLLFVOVRADPTLGGPHNTVT 600
DB 541 dRYVYHSLMGNAVDYLFIDESTGEIIFYSMDADFYHRONTLLFVOVRADPTLGGPHNTVT 600
QY 601 TOLVIELEDVNNTPPTLRRLRSTPVSVEENVPEGEYSREITATPDPTSAVLMFEIDMDST 660
DB 601 tolVIELEDVNNTPPTLRRLRSTPVSVEENVPEGEYSREITATPDPTSAVLMFEIDMDST 660
QY 661 WATKQREINTEYVGCIVETIYPTEGNRGSAIGRLVQVEIRDNTVIDEPEEMLYTLV 720
DB 661 watKQREINTEYVGCIVETIYPTEGNRGSAIGRLVQVEIRDNTVIDEPEEMLYTLV 720
QY 721 RYRLNLNVIGDDYEAFFETTTIIMNDNAPFANGTLTQMRVRELAAAGSLGVSATATD 780
DB 721 rYRLNLNVIGDDYEAFFETTTIIMNDNAPFANGTLTQMRVRELAAAGSLGVSATATD 780
QY 781 IDGPLYNGVRYTLQPRNTEPEGLKIDFTTGQIEVDNNEAIDAEPMRFLLYTVIASDE 840
DB 781 idgplYNGVRYTLQPRNTEPEGLKIDFTTGQIEVDNNEAIDAEPMRFLLYTVIASDE 840
QY 841 CSELENTECPDPSNYFEVPGDIEIEIIDTNKKVPEPLTEKNTTYVWENATSGDEVVOL 900
DB 841 cseLENTECPDPSNYFEVPGDIEIEIIDTNKKVPEPLTEKNTTYVWENATSGDEVVOL 900
QY 901 YSHDRDELHTVRYTMNFAVNPRLRDFEVDLDTGRLVYHVGDEKLRDGEPTHTI 960
DB 901 yshDRDELHTVRYTMNFAVNPRLRDFEVDLDTGRLVYHVGDEKLRDGEPTHTI 960
QY 961 FVNFIDNEFSDDGRNRDDEVEIEFVLLDVNDNAPEMLPDELRFVSEGAVALPE 1020
DB 961 fVNFIDNEFSDDGRNRDDEVEIEFVLLDVNDNAPEMLPDELRFVSEGAVALPE 1020
QY 1021 IYAPDRDEPTDNRKVGIGLIDLTITTDRIEVPDLFMISTENKTGELEFTAMDLRGWCT 1080
DB 1021 iYAPDRDEPTDNRKVGIGLIDLTITTDRIEVPDLFMISTENKTGELEFTAMDLRGWCT 1080
QY 1081 YEIIFEAHDGHPQORSNETYIVRYRPNFHPVFPQPSVIRLSERATEGVLATA 1140
DB 1081 yeIIFEAHDGHPQORSNETYIVRYRPNFHPVFPQPSVIRLSERATEGVLATA 1140
QY 1141 ANEFLEPIYATDEDLHAGSYTFHVQNEEAVQYFDITEVAGENSGLILRQLEPEQIR 1200
DB 1141 anEFLEPIYATDEDLHAGSYTFHVQNEEAVQYFDITEVAGENSGLILRQLEPEQIR 1200
QY 1201 QFRITIRATDGTGEPGLMTDVTESVVFVPIQGPVSEENATYVAFEGEGRLRSFLP 1260
DB 1201 qFRITIRATDGTGEPGLMTDVTESVVFVPIQGPVSEENATYVAFEGEGRLRSFLP 1260
QY 1261 QAEDELKMLCEBDQCDDIYRRFIDGNNEGLFVLDSSNYSIAQELDREAVSYTLHIAAS 1320
DB 1261 qAEDELKMLCEBDQCDDIYRRFIDGNNEGLFVLDSSNYSIAQELDREAVSYTLHIAAS 1320
QY 1321 NSPDATGPILOTSLVWVNVNREANRPRIEFQDLTYAGISTDUSIGRELLTVRASHTED 1380
DB 1321 nSPDATGPILOTSLVWVNVNREANRPRIEFQDLTYAGISTDUSIGRELLTVRASHTED 1380
QY 1381 TITTYTIDRASMLDSSLAVRDSAPALHATTGVLSLNQPTASMHGMEFDYIATDTASA 1440
DB 1381 tITTYTIDRASMLDSSLAVRDSAPALHATTGVLSLNQPTASMHGMEFDYIATDTASA 1440
QY 1441 IDTARVRYVLISSQNRVFTIPNQLETVYEQNRNFTAAFTSTCFNMTCNTDOVVPSSDSG 1500
DB 1441 idTARVRYVLISSQNRVFTIPNQLETVYEQNRNFTAAFTSTCFNMTCNTDOVVPSSDSG 1500
QY 1501 VADODTEVRAHFTBDNVQAOEVAVSDVLLRTIOLMSTNSLVLODVLVTDPTPL 1560
DB 1501 vADODTEVRAHFTBDNVQAOEVAVSDVLLRTIOLMSTNSLVLODVLVTDPTPL 1560
QY 1561 GEESMOIAVYALAAVAVGLCLVLLALFCRTRALNQLQALSMTRYKGSVDSGLNRAG 1620
DB 1561 geESMOIAVYALAAVAVGLCLVLLALFCRTRALNQLQALSMTRYKGSVDSGLNRAG 1620
QY 1621 LAPGINKHAVEGSSNPMWNEAIRAPFDALSDASGSDILGIDMQOFRDYFPDPDTS 1680

DB 1621 lappgcnkhavgesnmpmncatirapddaisdsgddliqiedmpqfrddyipppdtass 1680
 QY 1681 SGIVLHMGATDNKPYVTHGNNFGKSPRYLPQPHK 1717
 DB 1681 sgivlhmgatdnkpyvthgnnfgksprypypqphk 1717

RESULT 2
 AAB20499 standard; Protein; 1730 AA.
 AC AAB20499;
 XX 09-JUL-2001 (first entry)
 DE Corn earworm Bt toxin receptor.
 DE Corn earworm Bt toxin receptor.
 DE Corn earworm; CEM; Bacillus thuringiensis; Bt toxin;
 DE Receptor; crystal protein; CryIA; biological control; insecticide;
 DE Crop protection.
 XX Heliothis zea.
 XX
 XX Key Location/Qualifiers
 FH Binding-site 1305..1471
 FT /label=CryIA_binding_site
 FT Domain 1583..1601
 FT /label=Transmembrane_domain
 XX
 XX MO20013639-A2.
 XX 25-MAY-2001.
 XX 17-NOV-2000; 2000WO-US31674.
 XX 18-NOV-1999; 99US-0166285.
 XX 21-SEP-2000; 2000US-0234099.
 XX
 XX (PION- PIONEER HI-BRED ~~INC.~~ INC.
 XX Flannagan RD, Mathis JR, Meyer TE;
 XX MP1-2001-329223/34.
 XX N-PSDB; AAF30934.
 XX
 XX New nucleic acid encoding Bt toxin receptor, useful for screening and
 XX identifying new Bt toxin receptor ligands useful as new insecticidal
 XX toxins -
 XX
 XX Claim 4(a); Page 68-72; 85pp; English.
 XX
 XX The present sequence of is that of the Bacillus thuringiensis (Bt)
 XX toxin receptor of the corn earworm (CEW), Heliothis zea, which
 XX binds CryIA, particularly CryIA(b), toxins. The invention provides
 XX compositions and methods for modulating the susceptibility of a
 XX cell to Bt toxins. The compositions include Bt toxin receptor
 XX proteins from the Lepidopteran insects CEW, European corn borer
 XX (Ostrinia nubilalis) and fall armyworm (Spodoptera frugiperda).
 XX Nucleic acids encoding the polypeptides, antibodies specific to the
 XX polypeptides, and nucleic acid constructs for expressing the
 XX polypeptides in cells are also provided. The methods of the
 XX invention are useful for investigating the structure-function
 XX relationships of Bt toxin receptors, investigating toxin-receptor
 XX interactions, elucidating the mode of action of Bt toxins, screening
 XX and identifying novel Bt toxin receptor ligands including novel
 XX insecticidal toxins, and designing and developing novel Bt toxin
 XX receptor ligands. The methods are useful for managing Bt toxin
 XX resistance in plant pests, and protecting plants against damage by
 XX plant pests.
 XX
 XX Sequence 1730 AA;

Query Match 58.1%; Score 5179.5; DB 22; Length 1730;
 Best Local Similarity 59.5%; Pred. No. 0;
 Matches 1033; Conservative 240; Mismatches 426; Indels 37; Gaps 19;

QY 1 MGVE-RFPFAVLVLSAALANORCSYIATRPPEPELPIDVSGKWSQPLIPPT 59
 DB 1 mavdvrlitaavfliaahlta-qdcsymvaiprperpofpslndgylpwsrylplpveg 59
 QY REEVCMEENFLPDQM--IOYIYMEEEIEGDVITAKLNTYOGSNTP-VLSIMSGPRAQLPGE 116
 DB 60 redvcmeetqpalnprvcvifimeeelegdvatariingrntptlpsfsftgmipgv 119
 QY 117 FRQ-NEADGQMSLVITQRQDETATMOSYVSIOYEGSQAVLVALEIVNDNPIIQV 175
 DB 120 hrrlpengdwhlvltqgdyetpdmqyifdvrvedeplvatvmlivnddnpiilqm 179
 QY 176 VSACVIRPEHGEARLFDYOVSDRDGEISTRMPTFRVDSRADCSIFVMGEPSPDMF 235
 DB 180 fepcdipergelgtsckyyvdsadqelstrtmfeissdtdde-yfeivrenlqgwm 238
 QY 236 NMKMTVGINSPLNFETTLHFSVTASDLSLPNNHTVTMMVOYENESRPPKWEIYSVOQ 295
 DB 239 yvmrvtvkpkpdyeenplhlfrvayaslpthtvtmvgvenenpprmwelfavqg 298
 QY 236 FDEKTNOSFSLRAIDGDTGINRAIVTLIRDA--DFFSLYEIV--DGAILHVTIIRD 351
 DB 299 fdektnegsfvraldgdgldkpfyrlfetekegedlfsiqtleqregawfnvapird 358
 QY 352 KLERELFNLTIVAYKSTP-----ASFATPEAHIFITVDVNDQREPLHKESTIDIMET 405
 DB 359 tlekevinvsliaaykgynddvessfsqtcadvilivndvndqablpreeslsimeet 418
 QY 406 PWTLNFEFGFHRDLLENMAQYIVLEEDVPEPPGAASAFYIAPGSGYRQRTIMGTINHT 465
 DB 419 amtlnl-edfgfhdtrdlphagqyrvhlesinprahesfyiapvgyrqsgfimgtqh 477
 QY 466 MIDYEDVIFQNIIRKAVANDNASHVGEALVYVNLIMNDELPIFEESSTISASKETVIG 525
 DB 478 mldfvepfgnqlgravaidmdpdkwyaiainklilmndelpmfeedvqvsfdeleg 537
 QY 526 AGFPVATVLAIRDIDVYVSHMGNAVDYLFIDESTGEIFVSMDDADYRQNTLPFOV 585
 DB 538 agfyvatvavakrdvagdvevshmgnavsyrlidkevgellvteneahtyrhqealvqi 597
 QY 586 RADDTLGDGPHNTVTQLVIELEDVNNTPPTLRUPRSTPSVEENVEGYSREITATDP 645
 DB 598 paddtlge-pyntnttqvlvklrindnptlrlpratspsveenvpddgfpdqthardp 656
 QY 646 DTSATVAFEIMDSIMATKQGRETNPTVGCYIETIYPTBGNRGSAGLVQVQETIDN 705
 DB 657 dtlaelfeidwqnsyaskggrntdskeylgeletelylphndgrnaigvrvvrelidg 716
 QY 706 VTIDPEEFEMLYLTFRVYVNDLVNTIGDDVDEATFTTITIDMDNAPFANGTLTQRMVRE 765
 DB 717 vtiqtemfevlylvtivvndltvgedndistftittdmdndppliwegaltqefvire 776
 QY 766 LAASCTIGSVLAINDIDGLVYNQVRYTIQPRNNTPEGLVKIDFTTQGEIVANDNAIDA 825
 DB 777 vaasgvvlgsvlaiddidgplynyvryctltpldpedvldidntqgsvlqhaidede 836
 QY 826 PWRFLYTVTASDECSLENTKECPDSNFEVVEDIETIELITDNNKYPPELTKEFNITV 885
 DB 837 pprgnlyvtvasakcdllvtceppdpctyefetpgetihthdnkxpyveddkfeatv 896
 QY 886 YVMEVATGSDVWOLYSHDRDDELHYRTYTMFAVVRPRLDFFVEDLDGRLEVHPG 945
 DB 897 yiyegadgqgvvqiyasaeldrdeiyhkvsvyngayalnstrlrdffemldeaglyvmta 956
 QY 946 DEKLDRDGEDEPTHTITFVNFIQNFSDGGRNODVEIIFVLLDVNDNAPMPLPDELRF 1005
 DB 957 gelldrdgdgdepthtirfivndfnfyegdggrnqnetqylvllidndnype-lpetipw 1014

Oy	1066	DVSGNAGACRVATPELTPADPDROPPDNDNRNVCYGLIDLITDRDLFVSDLTFTMSIE---	1062
Db	1015	aisselgelervpepfactardcepgcdmsayavaltglastardqivpnllnmflliecdr	1074
Oy	1063	--NKTGELFEAMLRGCMGWGEYEFIEAFPHGAPDOORSNETVLVRPNFNHHPVFEPQ	1120
Db	1075	gidqltllaamdrlrywtgtygdldiaqdyhgipqrtsnqkypvlvrrprnfndrvtfvrpq	1134
Oy	1121	DSVIRLSREATEREGVLAATAANEEFLPVIATDDEGLHAQSVPFHVOGNEAVQYFDITEV	1180
Db	1135	gstlrlakeravnvgllatcvdgsgfldrlrtvaldedglaeglvftfsiaagddedagfdv--	1192
Oy	1181	GAGNSGQLILRLQLEPQRIKOFKITTRANDGSTEPEGPLMTDVTFESVVPYPOGDVPSEN	1240
Db	1193	ndgvnsqalltlrlfpfeeftefrvtlracdggltcpirsdclvctvrltpgtpevfedr	1252
Oy	1241	AATVAPEEGEGRRESPELPOAEDLNHLCEDDCCOITYRFIDGNNGELFVLDQSNYS	1300
Db	1253	tylvafrekeegmleeeelpiraadprnimceddchdtyslygnsgsehtfvdprrtnvs	1312
Oy	1301	LAEIIDEVATSYTLHLIAASNPDPATGIPLQTISILVTYNVEANRPRIPEODLYTAGIS	1360
Db	1313	lvkpldrsegethliligaedcnpbaav-1gasrtlvtvnvreanprpvfgralytagis	1371
Oy	1361	TLDISIGELLTVRASHTEDDTTYTIDRSKMOLDSELEVRBSALARTGVSLNMQP	1420
Db	1372	agdfierelltlvatshsediplrctligesmeadplleaavesafllmpetvslsnfcpr	1431
Oy	1421	TASHHGMEFEDVATDPAASIDPARVKVYLISSONVPTFPNOLETYEQNKNPFATRES	1480
Db	1432	tashmgfefevkatdsrtaretarkvyllsrnmvffitfmplpevcrgpedfaeetic	1491
Oy	1481	TGFNNNTNQVWPFS-D-SGVAKQDQTEVEAHFIKDNPVOAOEVEARSPDVTLLRQ	1539
Db	1492	afifmtcnidgswasaprvgaekdkdqterahfirddlpvpaeeeqrgprptlvnsiq	1551
Oy	1540	LMISTNSLVLODLVTGDTPTLGESSMOIAVAYALAASAVLGLCVLLIALFCRTALNR	1599
Db	1552	raleegnqlqiadlftgecptlllggdagaeralylaavaaalalilvvllliffvrrtlmr	1611
Oy	1600	QLOALSMTKYGVSQDSGLNRGL-APGTNRHAVGSNPMNEMEARPPDAISIDAQSDDL	1658
Db	1612	rlgalismtkyssqsgsinrvglaplnkhavgspimnetckapdfoalsesgydsdl	1671
Oy	1659	IGIDMPOFRDYPFGDPTDSSSGIV-LHMGEATDNKPPTTIGNNFGEKSPYLPQ	1713
Db	1672	igiedlpgfindyrpppegsgsmgvvnehpes-----lanhmnmfgfnspcfpspe	1722
Rd			
AA868331			
AA868331			
AC	AA868331		
XX	23-JUL-2004 (first-entry)		
XX	Pink bollworm-Bacillus thuringiensis toxin receptor Br-R2.		
XX	Pink bollworm-Bacillus thuringiensis, crystal toxin; CryIAc;		
KW	receptor; Br-R2; insecticide; pesticide; biological control;		
KW	Cotton.		
OS	Pectinophora gossypiella.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..20	
FT	Protein	/label= Signal_peptide 21..1729	
FT	Domain	/label= Mature_protein 1..1577	
FT		/note= "extracellular domain"	
FT	Domain	1578..1605	

FT		/note= "transmembrane domain"
FT	Domain	1606..1729
FT		/note= "cytoplasmic domain"
FT	Region	63..286
FT		/label= CR1
FT		/note= "putative cadherin motif 1"
FT	Region	287..293
FT		/label= CR2
FT		/note= "putative cadherin motif 2"
FT	Region	294..409
FT		/label= CR3
FT		/note= "putative cadherin motif 3"
FT	Region	410..536
FT		/label= CR4
FT		/note= "putative cadherin motif 4"
FT	Region	527..624
FT		/label= CR5
FT		/note= "putative cadherin motif 5"
FT	Region	625..770
FT		/label= CR6
FT		/note= "putative cadherin motif 6"
FT	Region	771..894
FT		/label= CR7
FT		/note= "putative cadherin motif 7"
FT	Region	895..1004
FT		/label= CR8
FT		/note= "putative cadherin motif 8"
FT	Region	1005..1154
FT		/label= CR9
FT		/note= "putative cadherin motif 9"
FT	Region	1155..1271
FT		/label= CR10
FT		/note= "putative cadherin motif 10"
FT	Region	1272..1365
FT		/label= CR11
FT		/note= "putative cadherin motif 11"
FT	Region	1366..1473
FT		/label= CR12
FT		/note= "putative cadherin motif 12"
FT	Modified-site	276
FT		/note= "N-glycosylated"
FT	Modified-site	380
FT		/note= "N-glycosylated"
FT	Modified-site	464
FT		/note= "N-glycosylated"
FT	Modified-site	500
FT		/note= "N-glycosylated"
FT	Modified-site	716
FT		/note= "N-glycosylated"
FT	Modified-site	769
FT		/note= "N-glycosylated"
FT	Modified-site	843
FT		/note= "N-glycosylated"
FT	Modified-site	893
FT		/note= "N-glycosylated"
FT	Modified-site	990
FT		/note= "N-glycosylated"
FT	Modified-site	1078
FT		/note= "N-glycosylated"
FT	Modified-site	1497
FT		/note= "N-glycosylated"
FT	Domain	1269..1367
FT		/note= "minimum binding fragment"
FT	Peptide	1355..1358
FT		/note= "READ signature sequence"
FT	Domain	1474..1577
FT		/note= "membrane proximal domain"
FT	Region	1541..1562
FT		/note= "leucine zipper motif"
FT	Region	1578..1600
FT		/note= "leucine zipper motif"
FT	Peptide	534..544
FT		/note= "unique peptide useful as antigen"

Query Match	Score	DB	Length
Best Local Similarity	55.38	DB 22:	Length 1729;
Matches 986; Conservativity 38.73;	Pred. No. 0;	Mismatches 422;	Indels 32; Gaps 18;
Sequence	1729 AA		
11 LTVLSASALANQ-----RCSTYII-AIPR-PEMPPELPIDYEKGSMEQOLPGCPREEYC	64		
11 VLTATVATSGGETSSCYMTDAIPRPKDDIPDLWTG-GWADWPIIPEPPDVC	69		
65 MENPLDQMIQ-----VIYMEETEGDVIATLKNYGSNTP--VLSJSGPRAQLAGE	116		
70 IINGWYP-QITSTSLGITIIHMEETEGDAIAKLNGDGSQTPREIVQPMWIGSNI-LSPE	127		
117 FRQNAADGQMSLVTRQDQYETATMQSYFSIQVGECAVLVALETIVINDNPLOAY	176		
128 Ir-ne-ngawlylietnrgdqetpimtryffdvtrpdetcaatvasislenidndpivryl	185		
177 SACTVPEHGEALPTCCYOVOSDRDEISRPMTFRDSSRADDESIFWYGVGEYSPSDMN	236		
186 DACQPELGEPIRLTCYQVSDGDRISLEPMTFLTSdr-edvqgIflyepahIgdwtn	244		
237 KMKITGINSPLLEFETQALHFEVSTASDSILPNHNTYTMQVENVESRPRAWETISVOQF	296		
245 mgItlgIIsalnfeenphIIfgtItaldsnpnhbtyvmwqyvenhIprwmeIfavqgf	304		
297 DEKTNQSFSLRAIDGDTGINRAINTYTLRDDADDFPSLEVI--EDGAILHYTEIDROKL	353		
305 demteqgfgvraIdtdgtgIkaIhyletdeeedIfIetclpgrnIdaIfstcamIdvdrI	364		
354 ERELNTLTVAKNSDASATFAHFIITVNDVNDORPELHKESYDIDMEENPMTINFE	413		
365 rTdvrIstIvwaKydnvsafrclprvIlyndInlnkKpqlIdgeyIstImeePrIsInfae	424		
414 EFGFHDRLGENAQVTELEDVFPFGAASAFYIAGSGQROQFTGINTHMLYEDVI	473		
425 IfgfYdeol-IyagslveIgegnprpvegaIfyIaptagqngftaIqgdhmlIdyedr	483		
474 FONIIKYKAVDMNASHVGEALVYVNLINMDELPIREESYSASFKETVGAGPRVATV	533		
484 fgnIkIklatdrclntfIgyaeavnvnlInmdeepIfeedqIvfkfketypkdyIvgsl	543		
534 LALRDIDDVVHSLMCNAVULYFIDESGELFEVMDAFLVHNQNTLVOYRADDTLSD	593		
544 rahddIdgsVvhsIlgnaantfIIdeeYgdIyuaIdaIdyIhrgneflIyraqdIme	603		
594 -GPHVTVTQVILEEDVNNPPTLRLPSPGSVEENVPEGEYSISRELTATPDTSAIYM	652		
604 pesrItlaaqIvleIdedvnnprcltrIprvspsveenvpegeIelnrelatpdrItayIq	663		
653 FEIDMDSTWATKQGETNPTEYVGCIVETIYPTGNSRGAISGLRVQVQEDRWATIDEE	712		
664 feIdwdtsafakqgdtnprIefhngvdtelIpnpadtreaugvIvakiInvnlIfhee	723		
713 FEMILYTRAVADLNVIGDDYDEAFITTIIDMNNAIPFANGITLTQMRAVELAASGL	772		
724 feIyIltvrvdIhIdddrYdesfIvIIdImdnwIwasgIInqIffsIstresstgv	783		
773 IGSVATPDIQPELVNQVYTTIOPRNNPBGVLIKDEFTGQIEVDNNAELIDDEPRFVLI	832		
784 IgsVatclIdpRIynqvYtIIPgedrYeqgIvqIIfvYrgqIIdvengaiIdprrwhI	843		
833 YTVIASDCSLSENRPCEPDSNYEVPDIDIEIITDNKKVPELPTEEFNNTVYWEANAT	892		

Dh	844	ytvlaasckseeneemscrrpbrvftvtdltidvnlvldimknkraadlsrfinetvyltlenpr	903
Oy	893	SGDEVLVYSHDRDRELtYHTFTVtMNFANVPRtRDFEYVLDITGtLEVNYPGDEKLDRD	952
Dh	904	dfenvvklysidedeidelytvtvqinayvnrldtffaidlsdsgvyeuennelldrd	963
Oy	953	GDEPTHTtPvNFIDNFESDGDGRNODENVtFVLLDNDNARFMRDLEtRtVDSGAV	101
Dh	964	rgeqhrhtfllnldfiseegdnrvnttevtvllldennarelpbtlepsviseedq	102
Oy	1013	AGVAVLDE--tYARDEDEPTDONSrVGSyILDLTRDRtE-VPRtFMtISENKGEL	106
Dh	1024	egtlldgesedvlyarpdldkedcptrshnvaylltmvltvrdtclvprllmndspnvtcgl	108
Oy	1069	ETtADtLlGtWGTtEIEtFEADtHtGtPOOtRSMEtTtVtLrRPFtNFtRtVtRPOtRtSLtR	112
Dh	1084	qtamrltgyvgtvlydslyatdhnlpqsthevelyeltrpnrhprpqtvfvrsegtllrld	114
Oy	1129	ERAtEGSVtLAtANEFLEtRYAtDEDGtLHAGSVtFtHvNGEAtVAYRtDLtENGAGENSQ	118
Dh	1144	eravnnvllslvngpdrldrtqaidddqldagvvtfdvlgadasnfyfn--ndgfsfgt	120
Oy	1189	LIILQLREPDtRQRtRtLrRtADtGtRGtRGtRtLMTDVTtSVtVtVtRQOtRtVtSENAAtVtARE	124
Dh	1202	llltqalpreeketevlratdgtterpsvtclstcltvtlvtprlsgpdlfgdntvsafte	126
Oy	1249	GEtEELtRESFELtPOtEDtKNtLCEtDCCDItYtRtIDtCNtEGLtFtLDDtSSNtVtSLtAELtRE	130
Dh	1262	keyglterfslphaedeknklctdcchdllyrtffgvydvrfdldrvtvllfkseldre	132
Oy	1309	VAItSYtLHtAASNSPDAtGtPLOTtSLtVtVtVtVtVtVtVtVtVtVtVtVtVtVtVtVtVt	136
Dh	1322	tthavgvaaasnptcgsglprlpgslltvtvtvredpvrfeqrltvgatstcdnlne	138
Oy	1369	LLtVtRtAhtEDDItTtYtIDtRtASQOLDSLEAVtDStAFALtATtGtVtSLtMNPtRtASHtGHtF	142
Dh	1382	llvtvratshenagqltyletdgsmaovdsltvaavdsarhlnagvavllltlqpraasmqmt	144
Oy	1429	EFVDtIADtTASAtIDtAVtKtVtLItSSORtKtVtFtDNOLEtEQRtNFtLAAtFtSGtFMtNtCN	148
Dh	1442	efnvatadpcektcdtaevkxyltssqnrtsffllndvetyesnrdfiaetfsvgfmmtcn	150
Oy	1489	IDOVtPESDSSGAOAddTtEVRAtHtFtRDNtVtVOAQtEAVtRSDtVtLrtTOLtMStNSLtv	154
Dh	1502	ldqyltptndagvlyqemaevnahfignlprvsadsleetrstqllttrsvgylngrrllv	156
Oy	1549	LQDLYtGDPtPLtGEtSQALVtALtAALtSAtVtGLtCLVtLLtALtECtRtRtALtRQOLtAStRK	160
Dh	1562	lndltvtpspdlgtgaqvtcltlylaqslsalfclllllltfvtvratlrrtltlealsmek	162
Oy	1609	YGSVtSGtNtRAGtL-APtCTNtHtAvtGtSPtMNPtNAtIRAPDPAItSDAtGSDtSLtIEtDMPQ	1666
Dh	1622	ygsavsgslntvgaiaapctknhalegsnrvpneqkacpdlasdzctsdssdllytdebslq	1680

	RESULT	4
AAW90182		
ID	AAW90182 standard; Protein; 1528 AA.	
XX	AC	
XX	AAW90182;	
DT	05-JUL-1999	(first entry)
XX		
DE	Manduca sexta BT toxin receptor BT-R	
XX		
KM	BT-R1; receptor; tobacco hornworm; F	
KW	BT-toxin; cadherin; insecticide; pes	
XX		
OS	Manduca sexta.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	I..21

CC	FT		/note= "signal peptide"
CC	FT	Protein	22..1528
CC	FT		/note= "mature protein"
CC	FT	Domain	1..1405
CC	FT	Domain	/note= "extracellular domain"
CC	FT	Domain	1406..1427
CC	FT	Domain	/note= "transmembrane domain"
CC	FT	Domain	1428..1528
CC	FT		/note= "intracellular domain"
CC	FT	Region	72..176
CC	FT	Region	/label= Cad1
CC	FT	Region	177..289
CC	FT	Region	/label= Cad2
CC	FT	Region	290..397
CC	FT	Region	/label= Cad3
CC	FT	Region	398..498
CC	FT	Region	/label= Cad4
CC	FT	Region	499..622
CC	FT	Region	/label= Cad5
CC	FT	Region	623..756
CC	FT	Region	/label= Cad6
CC	FT	Region	757..880
CC	FT	Region	/label= Cad7
CC	FT	Region	881..1004
CC	FT	Region	/label= Cad8
CC	FT	Region	1005..1121
CC	FT	Region	/label= Cad9
CC	FT	Region	1122..1241
CC	FT	Region	/label= Cad10
CC	FT	Region	1242..1352
CC	FT	Peptide	/label= Cad11
CC	FT	Peptide	470..500
CC	FT	Peptide	/note= "CNBr peptide"
CC	FT	Peptide	931..957
CC	FT	Peptide	/note= "CNBr peptide"
CC	FT	Peptide	1101..1119
CC	FT	Peptide	/note= "CNBr peptide"
CC	FT	Peptide	1257..1273
CC	FT	Peptide	/note= "CNBr peptide"
CC	XX		
CC	PN	WC9859048-A1.	
CC	XX		
CC	PD	30-DEC-1998.	
CC	XP		
CC	PF	08-JUN-1998:	98WC-US11868.
CC	XX		
CC	PR	20-JUN-1997:	97US-0880042.
CC	XX	(UYWY-) UNIV WYOMING.	
CC	PA		
CC	PI	Bulla LA:	
CC	XX		
CC	DR	WPI; 1999-105622/09.	
CC	DR	N-Psdb; AAV74277.	
CC	XX		
CC	PT	Newly isolated Bt-toxin receptor proteins	- useful for developing
CC	PT	improved pesticides	
CC	XX		
CC	PS	Claim 1; Fig 2; 6zpp; English.	
CC	XX		
CC	XX	The present sequence is Bt-R1, a glycoprotein receptor from the	
CC	CC	tobacco hornworm which binds to a Bacillus thuringiensis (Bt)	
CC	CC	toxin. Bt-R1 cDNA (see AAV75277) was obtained from a Manduca sexta	
CC	CC	midgut tissue cDNA library using degenerate probes based on isolated	
CC	CC	peptides (see AAW0183-87) of Bt-R1. Amino acid sequence analysis	
CC	CC	of the deduced protein shows that Bt-R1 is a member of the cadherin	
CC	CC	family, and has a highly repetitive external domain. The	
CC	CC	availability of Bt-R1 cDNA permits the retrieval of DNAs encoding	
CC	CC	homologous receptors in other insects and organisms as well as the	
CC	CC	design of assays for the cytotoxicity and binding affinity of	
CC	CC	potential pesticides and the development of methods to manipulate	
CC	CC	natural and/or introduced homologous receptors and, thus, to	
CC	CC	destroy target cells, tissues and/or organisms. The cDNA can also	

CC be utilised in the recombinant production of Br-R1 polypeptides
 CC that can be used: as targets for identifying agents that bind to
 CC Br-R1, as targets or bait for identifying and isolating binding
 CC partners and cellular factors that bind to Br-R1, as assay targets
 CC for identifying Br-R1 and other receptor-mediated activity, and as
 CC markers of cells that express a member of the Br-R1 protein family.
 CC They are also used in mediation of cell and tissue destruction,
 CC dissociation, dispersion, cell-to-cell association, and changes in
 CC morphology.

XX Sequence 1528 AA:

Query Match 52.7%; Score 4693; DB 20; Length 1528;
 Best Local Similarity 60.7%; Pred. No 0;
 Matches 925; Conservative 220; Mismatches 357; Indels 22; Gaps 13;

QY 1 MGVERFPVAVLVSLASALANORCYIIAIPR-PETPELPIDIDEGKSMSEQPLIPGP 59
 1 mavdvriaflvflapavlaqercymlprprdnlpvlnfegqtwsgqrpplape 60
 QY 60 REEVCMEF---LPDQIIVYIMEEIEGVYIAKLNGSNTPLVSI--MSGOPRAQLG 114
 61 rddicmdayhvitaniqltqlymdeeideftlaialnyngstpflelplfsgsynl-lm 119
 QY 115 PEHQNENADQMSLVITRQDDEYATQMSVVSIOVGEQAVLVLAETVINDNPRILQ 174
 120 pvltr---vdngsaahhaharqnyelpgmqymftrvqgsilvagvsalvlnldnapiq 177
 QY 175 VSAACVIPHEGEARLDCVYQVSDRDEISTRFMTFRVDSRRADESIFYAVGEYDPSDW 234
 178 nfeqrcvpejgeplrectcyqvadadgristefmtfidsvr-gdeetflyertnlpngw 236
 QY 235 FNMKMTVGINSPLEFETQIHFISVYASDSLPNNHYVTMMVQVENVESRPPKWEIIFSVQ 294
 237 mwlmntlqvtslnftvslphlftsvtaidslpnthvtmmvqvanvsrprvlelfavq 296
 QY 295 QPDEKTOQSLSRAIDDDTGINRAINTLIRDOADDFESLEVI---BDGALHVTIEDRD 351
 297 qfeeksyqntltvaidgdcltmplnrltlnedctfislealpbgsgsvflvpsldrd 356
 QY 352 KLERELFNLTIVAYKSTDSFATEAHFIIVNDVNDQRPPLRHEXSIDMEETPMPLNF 411
 357 tlgrevplrtivaykydeefststnvlvtindqrpplkeyrlalmeestpltnf 416
 QY 412 NEERGFHDRLGENAQYVLEEDVFPFGAASAFYIAGSGYOROTFMGTINHTMLDYED 471
 417 dkefgfndkdlgnaqytrlesvdpgraaefyiapevgvgrqtfimgltnhsmldyev 476
 QY 472 VIRQNIITIKKAVDMNNAHVGAALYVNLINNDLPRFEBSYSASFETVAGAFPA 531
 477 peltgsltrvvaalndltrgvallvldilnmdegpllehaqvcltldtegegfiva 536
 QY 532 TVLALDRIDIDVVVHSMGNAVLYLFIDESTGELFVSMDFAPYHQNQTFVQVRADDTL 591
 537 kavahndidvgvhehlllignavnfltdkltdgtirvsandsfnyheselfvyratdcl 596
 QY 592 GDGPHNTVTQVLEIEDVNNTPPLRLPRSTPSVEENVDEGEISRELTATPDTSAYL 651
 597 ge-pfhatsqlvrlindnlnrptlrlprsqpveenvpddhvlqelratpdtlacl 655
 QY 652 WPELIDMSTWATQGRRTNTEVVGCVIFETIYPTGKNRSALGRVLVDKIRNVITIDE 711
 656 rfeinwdsfatskqgrqanpdefrncveletlfealnrglaigrvareltlnhvidye 715
 QY 712 EFEMLYTVFVRDLNMYIGDYDEATFTIIRIDMNDNAIFPANGTTLQFVRVELAASGT 771
 716 efeslsltrvrdlnntvgyddgesmlltclimdnnapvwvgeltlegnfrvemsagql 775
 QY 772 LIGSVLATIDGPLYNQRVTTIOPRNTPEGVLAKIDFTTGQIEVDANEAIDAPWRFTY 831
 776 vvgsvradddgpllyngvrytlfpretdkdlmldlftglqisvntsgaldadtprfnl 835

QY 832 YTVIASDECSLENRRECPDPSNVFEPGDIEIEIIDTNKKEPELTERKENTVYWMENA 891
 836 ytvvaadrcstcpadcpdpdywtegnltlhdctmkyqaetcktdlvtvylayena 895
 QY 892 TSGDEVVQLXSHDRDELHYTRYTMNFAVNRBLRDFEVDLDGTGRLEHYNP--GDEK 948
 896 thldvvtliasldrdcldeyhytvsylnyvnprlmfmsvretglvydydctqsgsev 955
 QY 949 LDRDGEPTHTIFVNFIDNFSDDGRRNDEVEIFVLLDVNDNAPEMMLPDELFDVS 1008
 956 ldrdgepthtlfnldnmgsegrnqnqdevlilldvndnapelpppselstlws 1015
 QY 1009 EGAVAGVRVLPETIAPDRDPTDNRSGVGIIDLFTTRDIEVPLDFTMISLTKNTGEL 1068
 1016 enlkgyvrlphlfapdrdepdcnsvgyelnls-terdievpelfmqlanvtgel 1074
 QY 1069 ETAMDLRGVTEYELFEADHGYPOORSNETYLVIRPNFHPVFPVPPQPSVIRLSR 1128
 1075 etamdltkgyvtyalnlratdnglpqmsmetyelilhpfnypafvfrpndaavlrlar 1134
 QY 1129 ERATBEGVLAFAANFLEPIYATDEGLAAGSVTFHVQGNEEAVQYFDITTEGAGENSQ 1188
 1135 eravngvlatvngfelerlsatcdpqlhagvvtfgvvgdeesqrfyv--vndgenlgs 1192
 QY 1189 LILROLFPEQIRQFRITIRATDGTGEPGLMTDVTFSVFPVPGSPVSENAATVAFPE 1248
 1193 lrlilgvpelirefrlrltalcdqtdpgrlstcmctlrvtvfpqgeprlasselavafie 1252
 QY 1249 GEGELRESELPQAEPLDKNHLCEDDCODIYYRFDGNNGLFVLDDSSNVISLAQELDRE 1308
 1253 kasagmeesqplraqdikhnlceddchsllyrlidngseghfldpvrnrlflkkelire 1312
 QY 1309 VATSYLHLAASNSPRTATGPILOTSLVTVNVNREANPRPIFBDLYTIGISTIGRE 1368
 1313 qsashtlqvaaasnpdg-9lpilpasllvtvtrreadpvrfrlreltagistadsigre 1371
 QY 1369 LILVRASHTEPDITTYTIRASQWOLDSLEAVRDSAFALHATGVGLSLMMOPASWGMF 1428
 1372 lrlilhatqgsaalcyaldcltmvpsleavqsaefvlnaagvyltlnlpratahnglf 1431
 QY 1429 EPDVIATDPAASAIIDTARVVVYLIISSONRYTFIFDNOLETVENQNFIAATFTSGFMNTGN 1488
 1432 kfevtatdtaagadrdvrvyvsqnrvyfvrnltqgvdenrdtadtfaagfmntcn 1491
 QY 1489 IDQVPFSDS--SGVAODDTEVRA 1511
 1492 idqvpandpvtgvalehstqmaa 1515

RESULT 5
 AAB20500 standard; Protein; 1734 AA.
 AC AAB20500;
 DT 09-JUL-2001 (first entry)
 XX Fall armyworm Bt toxin receptor.
 DE Fall armyworm Bt toxin receptor.
 KW Fall armyworm; FAM: Bacillus thuringiensis; Bt toxin;
 KW receptor; crystal protein; CryIA; biological control; insecticide;
 KW crop protection.
 OS Spodoptera frugiperda.
 XX
 XX
 FH Key location/Qualifiers
 FT Peptide 1..22
 FT Protein /label-Signal_peptide
 FT 27..1734
 FT Binding-site /label-Mature_protein
 FT 1317..1486
 FT Domain /label-CryIA_binding_site
 FT 1594..1612

FT XX /label= Transmembrane_domain
 FN XX W0200136639-A2.
 PD XX 25-MAY-2001.
 XX
 PE 17-NOV-2000; 2000MO-US31674.
 XX
 PR 18-NOV-1999; 99US-0166285.
 PR 21-SEP-2000; 2000US-0234099.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Flannagan RD, Mathis JP, Meyer TE;
 DR WPI: 2001-3328223/34.
 DR N-PSDB: AAF30935.
 XX
 XX New nucleic acid encoding Bt toxin receptor, useful for screening and
 PT identifying new Bt toxin receptor ligands useful as new insecticidal
 PT toxins -
 PS
 PS Claim 4(a); Page 79-83; 85pp; English.

The present sequence of is that of the *Bacillus thuringiensis* (Bt) toxin receptor of the fall armyworm (FAW), *Spodoptera frugiperda*, which binds CryIA, particularly CryIA(b), toxins. The invention provides compositions and methods for modulating the susceptibility of a cell to Bt toxins. The compositions include Bt toxin receptor proteins from the lepidopteran insects FAW, European corn borer (*Ostrinia nubilalis*) and corn earworm (*Heliothis zea*). Nucleic acids encoding the polypeptides, antibodies specific to the polypeptides, and nucleic acid constructs for expressing the polypeptides in cells are also provided. The methods of the invention are useful for investigating the structure-function relationships of Bt toxin receptors, investigating toxin-receptor interactions, elucidating the mode of action of Bt toxins, screening and identifying novel Bt toxin receptor ligands including novel insecticidal toxins, and designing and developing novel Bt toxin receptor ligands. The methods are useful for managing Bt toxin resistance in plant pests, and protecting plants against damage by plant pests.

Sequence 1734 AA:

```

51.2%, Score 4561.5; DB 22; Length 1734;
Best Local Similarity 51.9%; Pred. No. 0;
Ches 906; Conservative 289; Mismatches 500; Indels 49; Gaps 23;

OY 1 MGVF-RFPFAVLVLSAALANQCSYIIAIPRETPELPIIDVGEKSMSPQPLIPPT 59
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 1 mavdrvlltatlvl-ltataqrdrccgymvelprdrpdrfpqndtqgltwagqpllpae 59

OY 60 REEVQENFNPDOMI-----QVIYEEETEGGVITAKINQGSNPVLSIMGQPRRA-OL 113
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 60 reevlndydepwsmnhngdqriymeeelgepvytlakinygnttpqrlpfrvgaahml 119

OY 114 GPFRPQ-NEADGQMSLVITTRQDVFETATMSTQVESIQVGEBSQAAVLVALEIVNDNPPT 172
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 120 gaeltreydatdwyllvltgrqdyetpdmgryltfqsvsgeglvltvtrldivndnapi 179

OY 173 LQVWACVPIPEHGEARLNDVCYQVSDRQGEISRTMTFRVDSRAADESIFPMWGEYDS 232
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 180 lemlepcnlpelvehvehckylvsdaagllstcsymssyndsder-gdevelrltkdypg 238

OY 233 DMFNKMTVGINSPNFEFTTOLHIPSVTASDSLPNNHIVTMVAVQVENESRPPRWVEIFS 292
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 239 dwtkywmllelksldygeophlfrfvtasdeljpnrlvtvmvvevnehrpmwelfa 298

OY 293 VQGFPEKRNQSSLSLAIDGDTGGINAINTLIRDDAD-DFESLEYI---EDGAILHYTEI 348
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 299 vqgfekqaksftvraiddglinpflrlyfletededefienlgegrdgarthvapi 358
```

QY	349	DROKLEREFNLTUVRKASND-----ASRPAHIEFTIVNUNOORPEPLKREXSIDM	402
QY	349	drdyLkrmfhtlriLaykgnndkgegesetsanvLlndndnqrrp ftkneyLism	418
QY	403	EETPEPTLNFNEEFORHBDJGNAQYVELLEDYPPGGAASAFIYAGSGYQROTIMGTI	462
Db	419	eteamtldl-qefghndrdignphayvnheslqpganlhaFyIaepgeyqsfstlgr	477
QY	463	NHTMLDYEDVLFQ-NIIIRKAYDMNNAHVGALYYVNLIMNNDELPIFEESYSASRQ	521
Db	478	lhmndyeddqyrgkiklkavaldrhnmhlgaaLhnlmndeplrfidedaynvte	537
QY	522	ETVGAFFAVNLALDIDIDVVVSHSLMGANVLYLTIDESTGEIYFSMDADVYRROWTL	581
Db	538	etvvgdfnlgryakdcldidvehsLlgnaenfltdldtdgyvsrddytdyqrneI	597
QY	562	FVQVRADDTLGDGPHNTVQTQVLIEEDVNNPPTJALPSPSPSEBNVBSYEIERSREIT	641
Db	598	lvqllavntllq-lpqnratcqltlfiedlnmtrpLlrrpsspyeeanvayhptlegit	656
QY	642	ATDPPTSAYLFEIEDMDSYATKOGRE-TNPTPYGCIYETIETPEGRKSGAISRLVWO	700
Db	657	atdpptcadiLfeIowmnsyatkqtrgnptadyhbcvealltyvpdpahygaeghlvar	716
QY	701	EIRDNVTLDEBEFEMLYLTVARVRLNLTVIGDDYDEATFTTIIIDMNDNPIFANCTLTOT	760
Db	717	evsddvltidyekfevLyLvrvtldmrvltpdydeamlvtvllidmdnwpIwadntllqct	776
QY	761	MVRRELAASGTLISVATIDIDGFLVQVAVYETIOPRNRPBEGLYKIDFTGGLIENDANA	820
Db	777	LrvremaagevLyvcllatdldpLjrnrvlymwpLkdrpddllaLnyvcqLtvnkgqa	836
QY	821	IDADEPMRFYLYTVVIASDECSL-ENRTECPDPSNYEYVGDIIEIIDIITNNKVEPLTE	879
Db	837	ldaddprfLylyvvtasdkcsldieffvcpdpqrCywntegelaIatldtmkIpraecd	896
QY	880	KENTYVYWMENATSGDEVOYLSHDRDRDELHRTVYTMNFAVNPRLRPFVDDTGL	939
Db	897	mtpsekriyentpugkIktllasdqdrdpnallYrlynaLhmlenlfavdpdltgeI	956
QY	940	EVHYHGDEKLDRODDEPTHTLFVNFIDNFSDDSGRNDODEEIVFVLLDNDVNAEMPL	999
Db	957	FvhtfsevlrdrdegepehrIftvlvnllegagdnqntlsevrvYllldndnkprrpI	1016
QY	1000	PDELEFDYSEBAVAGVRLPEITAPRDBDEPDTONSRVGGIIDLTTIDNDIEVP-DLPLM	1055
Db	1017	pdgefwltsvegevgkIrlpeIhndhdpndfndrsyvgIetsiklImndelapbdpfxI	1076
QY	1059	ISIEUK-----TGELEFAMLDRCGWGXYELFEAFDHFQOGRSNETTYLVRPNFNHP	1113
Db	1077	ltldldlctwkvgeletmtldrywgydvtelratldghfpmldsfecyglvtvrypnfshp	1133
QY	1114	VFVFPQDPDVIRLSRERATEGVLA--TANEFLEPIATDDEGLIAGSVTFHVQNEE	1170
Db	1137	vfvfptgsstlrIsteraIvngmlaIaIasgflarIstatedqIhagrvfIsIagnde	1196
QY	1171	AVQYDITIEVAGENSQOLIRQLRFPQIRQFRTIRARIDGCEBERPILMTDVTFSVNP	1230
Db	1197	aaeyfnv--lndgnaasmLtlkqalpbgvqgfvelvratatvggepgrIstsdsvvtvIm	1255
QY	1231	TQGDVFSSENAATYAFEGEGEGLKEESLFOAEDLNHLHCEDDCOQIYRFTIDGNNEGLF	1299
Db	1255	tqgdprfdidnaasvrtfyekegmeqfclpqaddpkrnyfcmddchIlystIvndgnghf	1314
QY	1291	VLDQSSNVISIAQELDREVATSYTLHATASNSPDATGIPJQTSILVYVYVWREANRPRTF	1355
Db	1315	avepelnvlyllkpldrdsqgeqyvvnvaasntbpgts-clssellvtIvgvreaanprpIf	1373
QY	1351	BODYLTATISLTDLSIGREGLTVARSHEDDTIYTTIDRASMSOLDSELAVRPSAFVLAHT	1410
Db	1374	eseIytavlnhtsdshelnylaakhsqIprIyvsldogemkldesIqtvcvetaIdlnsa	1433

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:08:22 ; Search time 70.14 Seconds
(without alignments)
1864.723 Million cell updates/sec

Title: US-09-715-909-2

Sequence: 1 MGVERFPFVLLVSLASAL.....THGNNGFPKSPVLPQHPK 1717

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sequences: 219241 seqs, 76174552 residues

TC: number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5607.5	62.9%	1715	JE0128	Bombyx mori receptor
2	602	6.8	5147	1 JEFFTM	cadherin-related t
3	560	6.3	4351	2 T00252	MEGFL protein - ra
4	516.5	5.8	3097	2 T00021	DN-cadherin - fru
5	498.5	5.6	2809	2 T30213	G-cadherin - sea u
6	492	5.5	1311	2 T33757	hypothetical prote
7	484.5	5.4	1439	2 T27110	hypothetical prote
8	452	5.1	3034	2 T14119	seven-pass transme
9	434.5	4.9	4307	2 T15276	hypothetical prote
10	405	4.5	2163	2 T120721	hypothetical prote
11	403	4.5	2610	2 T20968	hypothetical prote
12	374	4.2	3343	2 S44887	hypothetical prote
13	369.5	4.1	1865	2 S75200	fat protein - Syne
14	318	3.6	1069	2 T00043	BH-protocadherin-a
15	311	3.5	1180	2 T31066	vascular cadherin-
16	308.5	3.5	1072	2 T00041	BH-protocadherin-P
17	293	3.3	1200	2 T00042	BH-protocadherin P
18	291.5	3.3	882	1 IJHUCS	cadherin 1 precurs
19	290	3.3	1544	2 T29482	hypothetical prote
20	289.5	3.2	906	1 IJHUCN	N-cadherin precurs
21	283.5	3.2	889	2 T09055	protocadherin 68 -
22	279	3.1	785	2 T50180	cadherin-7 - chick
23	274	3.1	906	1 IJHUCN	cadherin 2 precurs
24	273	3.1	796	2 A38992	cadherin 11 precur
25	273	3.1	796	2 I48277	cadherin-11 - mous
26	271.5	3.0	790	2 T17016	N-cadherin-6 - huma
27	270.5	3.0	877	1 IJBOCN	N-cadherin precurs
28	269	3.0	701	2 T17243	hypothetical prote
29	269	3.0	796	2 A53584	OB-cadherin precur

30	267.5	3.0	790	2 I50178	cadherin-6B - chic
31	267	3.0	789	2 I52701	K-cadherin - rat
32	266	3.0	2205	2 T08615	aggregation factor
33	264.5	3.0	794	2 I59372	cadherin 12 - huma
34	264.5	3.0	796	2 I49556	cadherin-11 - huma
35	259.5	2.9	790	2 G02678	cadherin-14 - huma
36	258	2.9	790	2 I51638	F-cadherin - Afric
37	257	2.9	832	2 S55396	Li-cadherin - huma
38	251.5	2.8	871	2 S47518	cadherin - African
39	249.5	2.8	827	2 A53954	Li-cadherin precu
40	248.5	2.8	1329	2 T33136	hypothetical prote
41	243.5	2.7	906	1 IJXLC2	N-cadherin 2 precu
42	241	2.7	793	2 D38992	cadherin 8 - huma
43	238.5	2.7	783	2 I50116	N-cadherin precurs
44	236.5	2.7	912	1 IJHUCN	N-cadherin precurs
45	235	2.6	784	1 IJHUCS	cadherin 5 precurs

ALIGNMENTS

RESULT	1
JE0128	Bombyx mori receptor precursor - Bacillus thuringiensis
N:Accession:	Btr175
C:Species:	Bacillus thuringiensis
C:Date:	02-Jun-1998 #sequence-revision 10-Jul-1998 #text-change 07-May-1999
C:Accession:	JE0128
R:Nagatsatsu, Y.; Toda, S.; Koike, T.; Miyoshi, Y.; Shigematsu, S.; Kogure, M.	
Biosci. Biotechnol. Biochem. 62, 727-734, 1998	
A:Title:	Cloning, sequencing, and expression of the Bombyx mori receptor for Bacillus
A:Reference number:	JE0128; MUID:98276887
A:Accession:	JE0128
A:Molecule type:	mRNA
A:Residues:	1-1715 <NAG>
C:Keywords:	glycoprotein
F:1-21/Domain:	signal sequence #status predicted <SIG>
F:27-290/Domain:	proteoblast #status predicted <PR>
F:1567-1593/Domain:	transmembrane #status predicted <TM>
F:376-524	711,1069,1104,1640,1656/Binding site: carbohydrate (Asn) (covalent) #status

Query Match
Best Local Similarity 62.9% Score 5607.5; DB 2; Length 1715;
Matches 1095; Conserved 224; Pred. No. 0; Mismatches 377; Indels 29; Gaps 12;

QY	1	MGVERFPFVLLVSLASALANQRCYIIAIPRPETPELPIDYEGSKSWSEQLIPGPT	60
DB	1	MGVDVRLATLLILYAEVLAQERGFVAIPRPDLPEDFEGQTSQRLIPADR	60
QY	61	EYVCMENF--LPDMIOYIYMEEEIEGVYIYAKLNYQGSNTPVL--STMSCOPRAQLCP	115
DB	61	EYVCMENF--LPDMIOYIYMEEEIEGVYIYAKLNYQGSNTPVL--STMSCOPRAQLCP	115
QY	116	EPRNADQMSLVITRDYDFAFMQSVFSGVSGEASQVLALEIYINDNPPIIV	175
DB	121	IRRIIPDSNEMHILITRDYDFAFMQSVFSGVSGEASQVLALEIYINDNPPIIV	175
QY	176	VSACVLPENGEARLDCVYQVSDRDELSRPMTERVDSRRADSIFVGEYDPSDF	235
DB	181	LRPCQDELGEARLDCVYQVSDRDELSRPMTERVDSRRADSIFVGEYDPSDF	235
QY	236	NNKMTVGINSLPNETTOLHISVTSASDSLPNNHVTVMQVYENSRPRVVEIFSVQ	295
DB	240	RMTMTVGINSLPNETTOLHISVTSASDSLPNNHVTVMQVYENSRPRVVEIFSVQ	295
QY	296	EPEKTNOSFLAIDGPNINRAIVTLIRDDADDFESLEVE--DGAIIHVFTEIDBK	352
DB	300	EPEKTNOSFLAIDGPNINRAIVTLIRDDADDFESLEVE--DGAIIHVFTEIDBK	352
QY	353	LRRELFLNLTVAKSDASFAEAIHFIIVNDVNDORPEPLHREXSIDIMEPTPLNFN	412
DB	360	LRRELFLNLTVAKSDASFAEAIHFIIVNDVNDORPEPLHREXSIDIMEPTPLNFN	412

QY	413	EBEFGHROLDGNNQAYVLEVDYFPPGCAASATYIAPGSGYQQTITMGIINTHMLDYEDV	472
Db	420	QEFGFHRDLDGQNAQYVRLESDYPADKAAKFIAPVEVYQQTETIMGANHKMLDEYVP	479
QY	473	IFQOUIIKKAVADMMNASHGEALVYVNLINMDELPIFEESYSASFKEFYAGCFPAT	532
Db	480	EFGQIRLRVIATDMDNBEHGVAVYVINLINNMDEEPIEHSVQVNSPEFEGKCFEYAN	539
QY	533	VLAIDRIDVYVHSLMGNAVDYLFIDESTGELIYSMDADFTYHRONTLYVOVRADPTLG	592
Db	540	VRAIDRIDVREHETLMGANNMYLSIDKDTGDIHYQDFFPYHRSSELYVOVRADDTLG	599
QY	593	DGPHNYTQTOVILELEVONTPELTRPSTPEVEENVEGEISEITATPOTPSATYM	652
Db	600	E-PHRTATSQLIHLIEDINPTPTLPRGSPVEENVEEGIITISEIRADDPDTAELR	658
QY	653	FEIDMSTWATKQGRETNPTVEYVCVILEITYPEEGNBSAIGRLVQDEIRDNTIDEE	712
Db	659	FEIDMTSYATKQGREANPEFHNCVCEIETIYPAIINRSGAIGRLVVKIKRENTYIDEE	718
Q	713	FEMLYLVARVDLNTVYGGDYDEATFTTITIDNDNAPPIFANGTTLTOPKRVBELAOSTL	772
Db	719	FEMLYLVARRDLNTVYGGDYDESTFTTITIDNDNPPILWVGTLLEQSIRKREMSDAGV	778
QY	773	IGSVLATIDIGPLXNOVRYTIOBRNNTPEGLVKIDFTTGOILEVDAINEADAPERSFLY	832
Db	779	IGTLTADIDIGPLXNOVRYTKMKNCEGPENMLMIDYTGQIYVTKSGALDADVPERRNLY	838
QY	833	YTVIASOCSLENTPEPSPSNFEYVPGDLEIETIDTNKVEPELREKRNITYUYWEAT	892
Db	839	YTVIATDCAVEDDDCPPTWEPFGVQVYIETIDTNKIPOPEIDQKAVYVYIEYDAV	898
QY	893	SGDEVOLYLSHRRDDELKHTVRYTNMFANVRPLRDFEYVLDGTGRLEHYHGGDEKLDRD	952
Db	899	SGDEVVKVYIGSDLRDDIYHTIRQIYNVAPNRLDRFPAVDPTDGTGRVYVYTTDEVLDRD	958
QY	953	GDEPTHTIFVNFIDNPFSDODGRRNDQWEVIFVLLDVNDNAPEMPLDELRFVDSBQAV	1012
Db	959	GDEPOHRIFFELIDNPFQOGSDGNKNDQDAEVLVLLDVNDNAPELPEPELSWSVSESLT	1018
QY	1013	AGVAVLPEIYAPDRDEBDPTNSBVYGIJLDTTDRDIEVPLFTMISIENTKGLETFAM	1072
Db	1019	KGRFLQGHYIAPDRDEBDPTNSBVGAIIISLTIANEIEVPELFMIQIQTNGLETFAM	1078
QY	1073	DLRGWGTLEYLFEADFGHYGPQORSNETTYLVTRPNEHHVYVFPRPDPSYIRLSERAT	1132
Db	1079	DLRGWGTYVLIHIAVYDHGIRPQOMSNETEYLVTRPNFAHVPFPRKCATRLRLABERAV	1138
Q	1133	EGGVYLAAGNAEFEPYIYATQEDGILHGSGVTFHNGQEEAVOFFDTEVGAGNSQOLLR	1192
Db	1139	VNGLLATVDEGEFLKRIYATQEDGILHGQVAFEVYGTLEAVDYFHI--VNDGENSGTLMYK	1196
QY	1193	QLPEQIRQFRTIRATDGGTEGCPMLTDVTSVYFPTQSGDVPSENAATVAFPEGEG	1252
Db	1197	QLPEPDIRFEFVTRATDGGTEGTEPRPLSTCOTCFSEVYVPIQGEPIRPTSHYVAFIEKAG	1256
QY	1253	LRESFELPOAEDLKNHCEDDQODIYRFDGNGNELFVLDOSSNVISLAOLDBEVATS	1312
Db	1257	LLEHREHPRADRRKNHHCSDCHNIYRIIRIGDNDSDHEFLDETVAFLVYKLEDSVSET	1316
QY	1313	YTLRIASNSPDATGIPLOTSILVYVNVNAREAPRPIFODELYTAGISPTLDSIRELLTV	1372
Db	1317	YTLRIASNSPTG-GIAL--ISTITIYNNREADPOYFPRDITAGISISDSINNELLI	1374
QY	1373	RASHTEDDTTYTIDRASMOJSSLEAVRDSAFALHATTGYSILSMQPTASMHGFEEDV	1432
Db	1375	QATSEENAPRIYITDQMTWYTDPTPLASVRETAFLPHRGVLTILNOPSAMHGFEEOY	1434
QY	1433	IATDTASIDATKAKYVLISSONRVYFIENDQLEYVEOKRNFIAATFSGFMTCNDIYV	1492
Db	1435	VATPDAGISDPAANKYLLISITRRRVFELPNTLEOVQNTDIEIADTFSGAFMTCNDIYV	1494

Oy	1493	VPESSSSGVADDDTTEVNAHETIRNVNVPYQAEVEVRESDTVLATFTQLMTSTNSLVIQDL	1552
Oy	1493	: :: ::	:
Oy	1493	:: ::	:
Dd	1495	VPAIDASGAVINMGITEVGHFTRDWVPVADEIFELTKEDWMLALRAIOGTTLATRLVLVDL	1554
Oy	1553	VTGDPTTGEESSMQVAJVNALSAVLGFLVLVLIALFCFTRLNLROLQALSMTRYGSV	1612
Dd	1555	FT-DISPAFDAGSAAVLAVALSLALLALCLLLLYITIKTLNRLEALTYKKKGSV	1613
Oy	1613	DSGLNRACL-APGINKHAVEGSNPMMNEAIRAPDFDAISDASGSDLIIGIEDMPQFRDY	1671
Dd	1614	DSGLNRVIAIAPGTFKHAVEGSNPIMNETIKAPDFSMDASNSDILIGIEDLPFGENN	1673
Oy	1672	FPPGGTDSSSGVLHMGESATDNKP---VTTGNNNGEFTPYLDQ	1713
Dd	1674	YPRDVD-----EFKTDKPEDIVATRNNGEFTSPSPD	1708
RESULT 2			
JUEFTM			
cadherin-related tumor suppressor precursor - fruit fly (Drosophila melanogaster)	C:Species:	Drosophila melanogaster	
C:Date:	30-Sep-1993 #sequence_revision	30-Sep-1993 #text_change	16-Feb-1997
A:Accession:	A41087; B41087		
R:Mahoney, P.A.; Weber, U.; Onofrechuk, P.; Biesmann, H.; Bryant, P.J.; Goodman, C.S.	Cell 67,	853-868,	1991
A>Title:	The fat tumor suppressor gene in Drosophila encodes a novel member of the ca		
A:Reference number:	A41087; MUID:92069752		
A:Molecule type:	mRNA		
A:Residues:	143-485;1279-5147 <MAH>		
A:Cross-references:	GB:M80537		
A:Accession:	B41087		
A:Molecule type:	DNA		
A:Residues:	1-142;487-1278 <MA2>		
A:Cross-references:	GB:M80537		
A>Note:	1229-Gly and 1233-Ser were also found		
C:Genetics:			
A:Gene:	fat		
A:Cross-references:	FlyBase:FBgn0001075		
C:Superfamily:	cadherin-related tumor suppressor; cadherin repeat homology; EGF homol		
C:Keywords:	calcium binding; cell adhesion; duplication; transmembrane protein		
F:1-35/Domain:	signal sequence #status predicted <SIG>		
F:36-5147/Product:	cadherin-related tumor suppressor #status predicted <EXT>		
F:36-4583/Domain:	extracellular #status predicted <Ext>		
F:51-156/Domain:	cadherin repeat homology <CR1>		
F:159-270/Domain:	cadherin repeat homology <CR2>		
F:271-382/Domain:	cadherin repeat homology <CR3>		
F:390-494/Domain:	cadherin repeat homology <CR4>		
F:497-559/Domain:	cadherin repeat homology <CR5>		
F:602-708/Domain:	cadherin repeat homology <CR6>		
F:718-822/Domain:	cadherin repeat homology <CR7>		
F:831-942/Domain:	cadherin repeat homology <CR8>		
F:948-1049/Domain:	cadherin repeat homology <CR9>		
F:1052-1153/Domain:	cadherin repeat homology <CR10>		
F:1156-1278/Domain:	cadherin repeat homology <CR11>		
F:1281-1384/Domain:	cadherin repeat homology <CR12>		
F:1387-1489/Domain:	cadherin repeat homology <CR13>		
F:1492-1601/Domain:	cadherin repeat homology <CR14>		
F:1607-1713/Domain:	cadherin repeat homology <CR15>		
F:1717-1823/Domain:	cadherin repeat homology <CR16>		
F:1826-1922/Domain:	cadherin repeat homology <CR17>		
F:1925-2027/Domain:	cadherin repeat homology <CR18>		
F:2028-2167/Domain:	cadherin repeat homology <CR19>		
F:2169-2278/Domain:	cadherin repeat homology <CR20>		
F:2281-2384/Domain:	cadherin repeat homology <CR21>		
F:2387-2491/Domain:	cadherin repeat homology <CR22>		
F:2494-2596/Domain:	cadherin repeat homology <CR23>		
F:2599-2703/Domain:	cadherin repeat homology <CR24>		
F:2707-2810/Domain:	cadherin repeat homology <CR25>		
F:2813-2913/Domain:	cadherin repeat homology <CR26>		
F:2915-3013/Domain:	cadherin repeat homology <CR27>		
F:3014-3124/Domain:	cadherin repeat homology <CR28>		
F:3127-3229/Domain:	cadherin repeat homology <CR29>		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:19:57 : Search time 64.19 Seconds
(without alignments)
980.737 Million cell updates/sec

Title: US-09-715-909-2

Sequence: 1 MGVETFPFPAVLVSLASAL.....THGNNGFKSTPYLPQPHRK 1717

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	619	6.9	4590	1	FATH_HUMAN
2	594	6.7	5147	1	FAT_DROME
3	516.5	5.8	3097	1	CADN_DROME
4	374	4.2	3343	1	YOG7_CAEEL
5	339.5	3.8	789	1	CAD9_HUMAN
6	337.5	3.7	949	1	CDA8_HUMAN
7	325.5	3.5	1507	1	CAD8_DROME
8	314.5	3.5	950	1	CAD8_HUMAN
9	312.5	3.5	950	1	CAD8_HUMAN
10	311.5	3.5	798	1	CDBE_HUMAN
11	310	3.5	941	1	CDA4_HUMAN
12	308.5	3.5	941	1	CDA4_HUMAN
13	307	3.4	947	1	CDA4_HUMAN
14	307	3.4	1007	1	CDC2_HUMAN
15	307	3.4	1007	1	CDC2_HUMAN
16	303.5	3.4	788	1	CDA1_HUMAN
17	303	3.4	788	1	CDA1_HUMAN
18	301	3.4	818	1	CDB1_HUMAN
19	299	3.4	818	1	CDB1_HUMAN
20	297.5	3.3	930	1	CDA6_HUMAN
21	296	3.3	936	1	CDA5_HUMAN
22	295.5	3.3	797	1	CDB3_RAT
23	293	3.3	950	1	CDA9_HUMAN
24	291.5	3.3	797	1	CDB8_HUMAN
25	291.5	3.3	882	1	CDB8_HUMAN
26	290	3.3	776	1	CDBX_HUMAN
27	289.5	3.2	906	1	CAD2_MOUSE
28	289.5	3.2	930	1	CDA3_HUMAN
29	289	3.2	795	1	CDB4_HUMAN
30	288.5	3.2	801	1	CDB8_HUMAN
31	288	3.2	797	1	CDB9_HUMAN
32	286.5	3.2	795	1	CDBC_HUMAN
33	281	3.2	795	1	CDB5_HUMAN

34	278.5	3.1	906	1	CAD2_RAT	O91y3 rattus norv
35	277	3.1	800	1	CDBA_HUMAN	O9un67 homo sapien
36	275	3.1	793	1	CDB7_HUMAN	O9y5e2 homo sapien
37	274	3.1	906	1	CAD2_HUMAN	P19022 homo sapien
38	273	3.1	796	1	CAD8_HUMAN	P55287 homo sapien
39	271.5	3.0	790	1	CAD6_HUMAN	P55285 homo sapien
40	270.5	3.0	877	1	CAD7_BOVIN	P19534 bos taurus
41	270.5	3.0	937	1	CDA7_HUMAN	O9un72 homo sapien
42	269.5	3.0	792	1	CADB_CHICK	O93319 gallus gall
43	269	3.0	796	1	CADB_MOUSE	P55288 mus musculu
44	267.5	3.0	790	1	CAD6_CHICK	O90762 gallus gall
45	267	3.0	789	1	CAD6_RAT	P55280 rattus norv

ALIGNMENTS

RESULT 1	FATH_HUMAN	STANDARD	PRT	4590 AA.
ID	FATH_HUMAN			
AC	O14517			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	CADHERIN-RELATED TUMOR SUPPRESSOR HOMOLOG PRECURSOR (FAT PROTEIN HOMOLOG).			
GN	FAT.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphocytes;			
RX	MEDLINE=96163873; PubMed=8586420;			
RA	Dunne J., Hanby A.M., Poulson R., Jones T.A., Sheer D., Chin W.G.,			
RA	Da S.M., Zhao Q., Beverley P.C.L., Owen M.J.;			
RT	"Molecular cloning and tissue expression of FAT, the human homologue			
RT	of the Drosophila fat gene that is located on chromosome 4q34-q35 and			
RT	encodes a putative adhesion molecule.";			
RT	Genomics 30:207-223(1995).			
CC	- FUNCTION: COULD FUNCTION AS A CELL-ADHESION PROTEIN.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).			
CC	- TISSUE SPECIFICITY: EXPRESSED IN MANY EPITHELIAL AND SOME			
CC	ENDOTHELIAL AND SMOOTH MUSCLE CELLS.			
CC	- SIMILARITY: CONTRAINS 34 CADHERIN DOMAINS.			
CC	- SIMILARITY: CONTRAINS 5 EGF-LIKE DOMAINS.			
CC	- SIMILARITY: CONTRAINS 1 LAMININ G-LIKE DOMAIN.			
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	EMBL: X87241; CAA60685.1; -			
DR	HSSP: P01132; IEPH.			
DR	MIM: 600976; -			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR002126; Cadherin.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR000742; EGF-2.			
DR	InterPro: IPR001881; EGF-Ca.			
DR	InterPro: IPR001791; laminin_G.			
DR	pfam: PF00028; cadherin_33.			
DR	pfam: PF00008; EGF_5.			
DR	pfam: PF00054; laminin_G_1.			
DR	PRINTS: PR00205; CADHERIN.			
DR	SMART: SM00112; CA; 32.			
DR	SMART: SM00179; EGF_CA; 1.			
DR	SMART: SM00001; EGF_like; 5.			

